

Gibbs-Maße und Markov-Ketten

Von der Carl-Friedrich-Gauß-Fakultät
der Technischen Universität Carolo-Wilhelmina zu Braunschweig

zur Erlangung des Grades
Doktor der Naturwissenschaften (Dr. rer. nat.)

genehmigte Dissertation

von Tim Scharlau

geboren am 7. Juli 1979

in Bremen

Eingereicht am: 5. Mai 2008

Mündliche Prüfung am: 21. Oktober 2008

1. Gutachter: Prof. Dr. Thomas Sonar
2. Gutachter: Prof. Dr. Jens-Peter Kreiß
3. Gutachter: Prof. Dr. Andreas Meister

für meine Eltern und meinen Bruder

Zusammenfassung

Ende der sechziger Jahre wurde die Idee der Gibbs-Maße als stochastische Felder von Dobrushin, Lanford und Ruelle geprägt. Das Ziel war eine mathematisch rigorose Beschreibung der statistischen Mechanik. Im Laufe der Zeit haben sich die Gibbs-Maße neben der Bedeutung für die Physik als reiche Theorie im Bereich der Stochastik erwiesen. So sind z. B. unter bestimmten Voraussetzungen Gibbs-Maße äquivalent zu Markov-Ketten. Ziel meiner Arbeit ist es den Begriff des Gibbs-Maßes verständlich einzuführen und erstmalig unterschiedliche Aspekte der Gibbs-Maße von der Definition als stochastisches Feld über stationäre Maße von Prozessen und zellularen Automaten bis zur Äquivalenz zu Markov-Ketten zusammenzustellen. Das ist in dieser Form und Themenauswahl bisher nicht geschehen, so dass Anfänger eine lange Einarbeitungszeit in das Gebiet der Gibbs-Maße benötigten und die Zusammenhänge zu anderen Bereichen der Mathematik verloren.

Als besonders fruchtbar stellt sich die oben erwähnte Äquivalenz von Gibbs-Maßen und Markov-Ketten heraus. Dies zeigt sich z. B. in der Verwendung von Gibbs-Maßen bei der Analyse von DNS.

Da ich selbst das erste Staatsexamen für das höhere Lehramt in Mathematik und Physik gemacht habe und in Zukunft als Lehrer an Gymnasien arbeiten werde, schließt die Arbeit mit einem Kapitel zur Rolle von Markov-Ketten in der gymnasialen Oberstufe. Neben der Verbindung von Analysis, Lineare Algebra und Stochastik in der Oberstufe, bietet sich u. a. durch die unter bestimmten Voraussetzungen vorliegende Äquivalenz von Markov-Ketten und Gibbs-Maßen die Möglichkeit, fächerübergreifend zu unterrichten. Denn fasst man die Parametermenge der Markov-Kette als eine räumliche statt zeitliche Menge auf, so kommt man z. B. zum Ising-Modell.

Abstract

The idea of Gibbs measures as random fields was developed by Dobrushin, Lanford and Ruelle at the end of the 1960s. They aimed at a rigorous mathematical description of statistical mechanics. In the course of time Gibbs measures have proved to be valuable not only in physics but also in probability theory. Thus, under certain circumstances Gibbs measures are equivalent to Markov chains. In my thesis, I want to introduce the term Gibbs measures in an understandable way and bring together different aspects of Gibbs measures - from their definition as a random field to stationary measures of processes and cellular automata or the equivalence to Markov chains. This has not been done in this extent before so that it has taken beginners a lot of time to make themselves acquainted with the field of Gibbs measures. Moreover, I intend to connect the theory of Gibbs measures with other mathematical fields.

The equivalence between Gibbs measures and Markov chains turned out to be especially fruitful. As a consequence, Gibbs measures are used for DNA analyses.

Since I am going to be a grammar school teacher for mathematics and physics, my thesis ends with a chapter about the use of Markov chains in school. Apart from the connection between analysis, linear algebra and probability theory, the equivalence of Gibbs measures and Markov chains, which is given under certain circumstances as mentioned above, offers an opportunity to integrate several school subjects: If the parameter set of a Markov chain is taken as a spatial set instead of a temporal set we can get to the Ising model.

Danksagungen

Diese Arbeit ist im Rahmen meiner Tätigkeit als wissenschaftlicher Mitarbeiter an der Technischen Universität Braunschweig im Institut *Computational Mathematics*, AG Partielle Differentialgleichungen entstanden. Ich möchte meinem Mentor Prof. Dr. Thomas Sonar sehr herzlich für diese großartige Möglichkeit und die weit über die fachliche Betreuung hinausgehende Unterstützung in meiner Assistenten- und Studienzeit, also in den letzten beinahe acht Jahren, danken.

Ich danke allen Mitgliedern der AG Partielle Differentialgleichungen für die Unterstützung, tolle Zusammenarbeit und die wunderbare Zeit im Institut. Insbesondere danke ich Oliver Nowak für die Durchsicht und Korrektur einer ersten Version dieser Arbeit.

Inhaltsverzeichnis

1. Einführung	1
I. Gibbs-Maße	7
2. DLR-Gleichung	9
2.1. Grundlagen	9
2.2. Wahrscheinlichkeitskerne und Spezifizierungen	11
2.3. Gibbssche Spezifizierung	13
2.4. Existenz und Phasenübergänge	18
3. Markov-Prozeß und zellulare Automaten	25
3.1. Stochastischer Prozeß und Markovsche Halbgruppe	25
3.2. Markov-Prozeß	27
3.3. Übergangswahrscheinlichkeiten kontinuierlicher Markov-Ketten	29
3.4. Stationäre Maße kontinuierlicher Markov-Ketten	33
3.5. Bedingte Gibbs-Maße als stationäre Maße einer kontinuierlichen Markov-Kette	35
3.6. PCA und ESM	38
3.7. Beispiel: Vom PCA zum ESM	40
3.8. Eine Klasse von zellularen Automaten	42
II. Markov-Felder und Markov-Ketten	45
4. Zusammenhänge von Markov-Feldern, Markov-Ketten und Gibbs-Maßen	47
4.1. Ergodensatz für Markov-Ketten	47
4.2. Stochastische Felder	49
4.3. Homogene Markov-Spezifizierungen und Gibbssche Spezifizierungen	51
4.4. Das eindimensionale Ising-Modell	58
4.5. Sequenzen von Nukleotiden	61
5. Markov-Ketten in der Schule	67
5.1. Einleitung	67
5.2. Der diskrete Putzer-Algorithmus	69
5.3. Beispiele	70

1. Einführung

Ich beginne mit der physikalischen Idee von Gibbs-Maßen. Will man Flüssigkeiten oder Gase mikroskopisch beschreiben, so steht man vor einem großen Problem, nämlich der riesigen Anzahl von Molekülen. Das betrachtete System ist also außerordentlich komplex. Selbst wenn wir die das System beschreibende Differentialgleichung, die Newtonsche Bewegungsgleichung bzw. die Schrödingergleichung, lösen könnten, hätten wir das Problem, den Zustand zu einem späteren Zeitpunkt zu bestimmen, da wir eine nicht zu beschaffende Anzahl von Anfangsbedingungen benötigten. Da jedoch für die makroskopische Beschreibung wenige Mittelwerte, die makroskopischen Größen, ausreichen, ist der Ansatz der statistischen Physik die Beschreibung des Systems mit Wahrscheinlichkeitsmaßen. Aber auch dieser Zugang birgt viele Probleme. Um die Schwierigkeiten ein wenig zu begrenzen, werde ich Spinsysteme betrachten, d. h. die Teilchen können sich nur auf einem Gitter bewegen.

Häufig werden als weitere Vereinfachungen endliche Systeme betrachtet. Dort hat man eine endliche Teilmenge $\Lambda \subset \mathbb{Z}^d$ als Gitter und jeder Gitterpunkt $i \in \Lambda$ kann besetzt oder unbesetzt sein, also die Zustände 1 oder 0 annehmen. Die Elementarereignisse ω werden Konfigurationen genannt und sind Elemente aus $\Omega = E^\Lambda$ mit $E = \{0, 1\}$. Den Spin an der Stelle $i \in \Lambda$ bezeichnet man mit $\omega_i \in \{0, 1\}$. Jeder Konfiguration wird mit der Hamilton-Funktion $H : \Omega \rightarrow \mathbb{R}$ eine Energie zugeordnet. Ist $\beta = \frac{1}{kT}$, wobei T die absolute Temperatur und k die Boltzmann-Konstante ist, so wird durch

$$p : \Omega \rightarrow [0, 1], \quad p(\omega) = \frac{\exp(-\beta H(\omega))}{Z} \quad (1.1)$$

ein Wahrscheinlichkeitsmaß auf Ω definiert, das als endliches Gibbs-Maß bzw. großkanonisches Ensemble bezeichnet wird. H beinhaltet das sogenannte chemische Potential h , welches beispielsweise die benötigte Energie angibt, um ein Teilchen im System zu platzieren. Z ist die Normierungskonstante bzw. Zustandssumme mit

$$Z := \sum_{\omega \in \Omega} \exp(-\beta H(\omega)).$$

Das endliche Gibbs-Maß wird zur Beschreibung von Systemen im Gleichgewicht genutzt. Ein System befindet sich im Gleichgewicht, wenn sich das das System beschreibende Wahrscheinlichkeitsmaß mit der Zeit nicht mehr ändert und sich damit die makroskopischen Größen, die mit Hilfe der Wahrscheinlichkeitsmaße berechnet werden, ebenfalls nicht mehr verändern. Warum das endliche Gibbs-Maß einen Gleichgewichtszustand beschreibt ist nicht rigoros geklärt. Eine Begründung benötigt viele Postulate. Deshalb kann man auch an dieser Stelle postulieren, dass (1.1) den Gleichgewichtszustand beschreibt, vgl. [27], [31], [53], [63], [72], [73]. Dennoch möchte ich kurz auf die Begründung

1. Einführung

der endlichen Gibbs-Maße eingehen. Der erste Schritt ist die Einführung von statistischen Ensembles. Man denke an ein System bestehend aus vielen Teilchen, das einen bestimmten Makrozustand besitzt. Mikroskopisch gibt es viele mögliche Realisierungen des Systems, die mit der Zeit angenommen werden können. Jede Realisierung unterscheidet sich jedoch durch die Konfiguration und Energie. Die Gesamtheit aller Realisierungen wird als statistisches Ensemble bezeichnet. Ein Ensemble besteht also aus vielen unabhängigen Systemen. Im Fall der konstanten Temperatur, des konstanten Volumens und des konstanten chemischen Potentials spricht man vom großkanonischen Ensemble. Hier können die Energie und die Teilchenzahl variieren. Nun benötigen wir das erste Postulat, die Quasiergodenhypothese, wonach die tatsächliche Realisierung des Systems allen möglichen Realisierungen beliebig nahe kommt. Diese Voraussetzung können wir nutzen. Eine von der Konfiguration ω abhängige Observable F ist ebenfalls zeitabhängig, da sich die Konfiguration mit der Zeit verändert. Wollen wir einen Mittelwert \bar{F} der Observablen berechnen benötigen wir die Quasiergodenhypothese und eine unendliche Beobachtungsdauer, um die Abhängigkeit von den unbekannten Anfangsbedingungen zu eliminieren

$$\bar{F} = \lim_{t \rightarrow \infty} \int_0^t F(\omega(t)) dt.$$

Nun bekommen wir aber die zu Beginn angesprochenen Probleme mit der Berechnung der Konfiguration zu einer späteren Zeit, ω kann im Allgemeinen nicht explizit in Abhängigkeit von der Zeit angegeben werden. Dieses Problem lösen wir mit dem großkanonischen Ensemble. Das endliche Gibbs-Maß gibt die Wahrscheinlichkeiten für alle möglichen Konfigurationen an und wir können wieder mit der Quasiergodenhypothese das Zeitmittel durch ein Scharmittel $\langle F \rangle$, definiert durch

$$\langle F \rangle = \sum_{\omega} F(\omega) p(\omega),$$

ersetzen. Kurz gesagt fordern wir mit der Quasiergodenhypothese

$$\bar{F} = \langle F \rangle.$$

Dies klärt aber noch nicht die Herkunft des endlichen Gibbs-Maßes, sondern die Anwendung. Für eine Andeutung der Herkunft brauchen wir neben der Quasiergodenhypothese die für die statistische Physik grundlegende Annahme, dass alle mit einem gegebenen Makrozustand verträglichen Mikrozustände mit der gleichen Wahrscheinlichkeit auftreten. Die Herleitung von (1.1) geht nun von einem offenen System mit konstanter Temperatur, konstantem chemischen Potential und konstantem Volumen aus. Das System ist in ein großes Teilchen- und Wärmereservoir eingebettet und bildet zusammen mit diesem Reservoir ein isoliertes System. Das isolierte Gesamtsystem hat also eine konstante Energie E und Teilchenzahl N . Befindet sich das betrachtete offene System in einem Mikrozustand i mit Teilchenzahl N_S und Energie E_S , so bleiben für das Reservoir noch eine große Anzahl Γ von Zuständen mit der Teilchenzahl $N - N_S$ und der Energie $E - E_S$. Die Wahrscheinlichkeit p_i für den Mikrozustand i soll nun proportional zu Γ sein. Das Resultat dieser Überlegungen liefert (1.1), siehe [31], [53], [72], [73].

Hält man die Teilchenzahl konstant, so gelangt man zum kanonischen Ensemble, für welches h aus der Hamilton-Funktion verschwindet, siehe [25]. Für unsere mathematischen Betrachtung spielen β und h keine Rolle bzw. können sie, wie für das h bereits geschehen, in die Hamilton-Funktion eingearbeitet werden. Ich werde die Teilchenzahl nicht fixieren, so dass die Hamilton-Funktion h und β beinhaltet, wie z. B.

$$H(\omega) = - \sum_{\{i,j\} \subset S} J(i,j) \omega_i \omega_j - h \sum_{i \in S} \omega_i. \quad (1.2)$$

Es handelt sich also um das großkanonische Ensemble im Fall der Gittergase. Dennoch möchte ich an dieser Stelle das kanonische Ensemble nutzen, um eine weitere Herleitung von (1.1) zu liefern. Wir betrachten also ein geschlossenes System im Wärmebad. Nur die Energie variiert. Diese von E.T. Jaynes aus der Informationstheorie stammende Herleitung wird als das „Prinzip der maximalen Entropie“ bezeichnet, siehe [4], [36], [76]. Für eine detaillierte Charakterisierung von Gibbs-Maßen durch die Eigenschaft der maximalen Entropie verweise ich auf das Kapitel zum Variationsprinzip in [26] und [23]. Wir werden am Ende der Herleitung die Gleichungen aus der Thermodynamik wieder erkennen. Zunächst benötigen wir die Definition der Shannon-Entropie S eines Wahrscheinlichkeitsmaßes μ auf Ω

$$S(\mu) = - \sum_{\omega \in \Omega} \mu(\omega) \ln \mu(\omega), \quad (1.3)$$

wobei $0 \ln 0 := 0$. Die Shannon-Entropie kann als Maß der Unsicherheit interpretiert werden. Das Prinzip der maximalen Entropie arbeitet mit der Methode von Lagrange, d. h. wir maximieren (1.3) unter bestimmten Nebenbedingungen, vgl. [36] Kapitel 11.6. Im Fall des kanonischen Ensembles haben wir zwei Nebenbedingungen. Als erste Nebenbedingung erhält man

$$\sum_{\omega \in \Omega} \mu(\omega) = 1. \quad (1.4)$$

Die Energie ist zwar nicht konstant, dennoch ergibt sich im Gleichgewicht ein konstanter Mittelwert $\langle H \rangle$ mit

$$\langle H \rangle = \sum_{\omega \in \Omega} \mu(\omega) H(\omega) \quad (1.5)$$

und damit die zweite Nebenbedingung. Um die Schreibweise etwas zu vereinfachen, setze ich die Anzahl der Elemente von Ω gleich n und schreibe statt $\sum_{\omega \in \Omega} \mu(\omega) = 1$ für die Herleitung

$$\sum_{i=1}^n p_i = 1$$

und analog

$$\sum_{i=1}^n p_i H_i = 1.$$

1. Einführung

Wir können jetzt für die Maximierung von (1.3) unter den Nebenbedingungen (1.4) und (1.5) die Lagrange-Funktion aufstellen

$$L(p_1, \dots, p_n, \lambda_1, \lambda_2) = - \sum_{i=1}^n p_i \ln p_i + \lambda_1 \left(\sum_{i=1}^n p_i - 1 \right) + \lambda_2 \left(\sum_{i=1}^n p_i H_i - \langle H \rangle \right).$$

Bilden wir $\nabla L = 0$, so erhalten wir die beiden Nebenbedingungen und

$$\frac{\partial L}{\partial p_i} = -(\ln p_i + 1) + \lambda_1 + \lambda_2 H_i = 0.$$

Wir bekommen

$$p_i = \exp((\lambda_1 - 1) + \lambda_2 H_i).$$

Aus (1.4) folgt

$$\lambda_1 = - \sum_{i=1}^n \lambda_2 H_i + 1$$

und damit

$$p_i = \frac{1}{\sum_{i=1}^n \lambda_2 H_i} \exp(\lambda_2 H_i), \quad (1.6)$$

was schon unserem kanonischen Ensemble entspricht. Mit (1.5) folgt λ_2 .

Nutzen wir die Hauptsätze der Thermodynamik, ebenfalls Postulate, bekommen wir für ein geschlossenes System im Wärmebad die Bedingung, dass die freie Energie F im Gleichgewichtszustand minimal ist, wobei

$$F = \langle H \rangle - TS_{Thermo}$$

und sich die Entropie der Thermodynamik S_{Thermo} aus der Shannon-Entropie S mit $S_{Thermo} = kS$ ergibt, siehe [4], [52], [53]. Ebenfalls aus der Thermodynamik wissen wir, dass

$$F = -kT \ln \left(\sum_{i=1}^n \exp(\beta H_i) \right) = -kT \ln Z$$

bzw.

$$Z = \exp(-\beta F).$$

Setzen wir mit diesem Wissen in (1.6) $\lambda_2 = \beta$, so folgt

$$p_i = \exp(\beta(F - H_i)).$$

Bildet man auf beiden Seiten den Logarithmus, erhalten wir

$$\ln p_i = \beta(F - H_i)$$

und damit

$$\sum_{i=1}^n p_i \beta(F - H_i) = \beta(F - \langle H \rangle).$$

Nun gilt aber nach Definition (1.3)

$$S = \beta(\langle H \rangle - F) = \frac{1}{kT}(\langle H \rangle - F)$$

und wir erhalten für die freie Energie die Gleichung aus der Thermodynamik

$$F = \langle H \rangle - TS_{Thermo}.$$

Die Beschränkung auf ein endliches Gitter bringt einen Nachteil mit sich, Phasenübergänge können mit dem endlichen Gibbs-Maß nicht beschrieben werden. Betrachtet man beispielsweise Wasser, so müßte die Dichte in Abhängigkeit der Temperatur Unstetigkeiten aufweisen, was mit dem endlichen Gibbs-Maß für Temperaturen $T > 0$ nicht zu erklären ist. Das ist ein Grund, warum die Betrachtung von unendlichen Gittern, z. B. $S = \mathbb{Z}^d$, eine große Rolle spielt. Läßt man unendliche Gitter zu, entsteht ein neues Problem, die Definition (1.1) des Gibbs-Maßes kann auf $\Omega = E^S$ nicht aufrechterhalten werden. Die Ursache dafür ist die Definition der Hamilton-Funktion H , siehe Kapitel 2. Statt Wahrscheinlichkeitsmaßen werden nun Familien von bedingten Wahrscheinlichkeiten, die so genannten Spezifizierungen, studiert. Anders als in der Stochastik, wo ein Wahrscheinlichkeitsmaß gegeben ist und bedingte Wahrscheinlichkeiten damit berechnet werden, sucht die statistische Physik Wahrscheinlichkeitsmaße, deren bedingte Wahrscheinlichkeiten durch die Spezifizierung gegeben sind. Ich will in dieser Arbeit eine besondere Spezifizierung, die Gibbssche Spezifizierung, betrachten. Wahrscheinlichkeitsmaße auf $\Omega = E^S$, deren bedingte Wahrscheinlichkeiten mit denen der Gibbsschen Spezifizierung übereinstimmen, werden Gibbs-Maße genannt. Diese Maße beschreiben wieder, wie die endlichen Gibbs-Maße, Gleichgewichtszustände. Erfüllen die bedingten Wahrscheinlichkeiten von mehr als einem Maß die Gibbssche Spezifizierung, so kann das System zwischen mehreren Gleichgewichtszuständen wählen. Dieses Verhalten wird als Phasenübergang interpretiert.

In Kapitel 2 führe ich die Familie von bedingten Wahrscheinlichkeiten, genauer gesagt die Familie von Wahrscheinlichkeitskernen, ein, die die Gibbssche Spezifizierung bildet. Damit bekommen wir wie oben erläutert, die Definition der Gibbs-Maße. Dass es tatsächlich zumindest ein Maß gibt, dass der Gibbsschen Spezifizierung genügt ist keinesfalls selbstverständlich und muss sogar in einigen Fällen verneint werden. Im letzten Teil des Kapitels 2 führe ich für bestimmte Voraussetzungen einen Existenzbeweis.

Die Definition der Gibbs-Maße mit Hilfe der Familien von Wahrscheinlichkeitskernen erinnert an die Konstruktion von Markov Prozessen. Dies wird durch die an stochastische Prozesse angelehnte Definition von stochastischen Feldern deutlich. So wird das Quadrupel $(\Omega, \mathfrak{F}, \mu, (\pi_i)_{i \in S})$ als stochastisches Feld bezeichnet. Dabei sei $(\pi_i)_{i \in S}$ eine Familie von Zufallsvariablen mit dem gemeinsamen Zustandsraum (E, \mathfrak{E}) , S sei die Parametermenge und $(\Omega, \mathfrak{F}, \mu)$ ein Wahrscheinlichkeitsraum. Ist μ ein Gibbs-Maß, so wird das Feld als Gibbs-Feld bezeichnet. Die Parametermenge S ist anschaulich das Gitter und der Zustandsraum $E = \{0, 1\}$ enthält die Zustände unbesetzt und besetzt. Betrachtet man statt eines Gittergases z. B. das mathematische Modell eines Magneten, das Ising-Modell, wird der Zustandsraum zu $E = \{-1, +1\}$, „spin down“ oder „spin up“.

1. Einführung

Die Rolle der Gibbsschen Spezifizierung spielt in Kapitel 3 die Markovsche Halbgruppe, ebenfalls eine Familie von Wahrscheinlichkeitskernen. Diese Familie definiert allerdings kein Feld sondern einen Prozeß. Der Prozeß wird als Quadrupel $(\Omega, \mathfrak{F}, \mu, (X_t)_{t \in T})$ definiert. Die Parametermenge T der Familie von Zufallsvariablen ist allerdings eindimensional, $T \subset \mathbb{R}$. Dieser Unterschied der Indizierung wird schon bei den Familien von Wahrscheinlichkeitskernen deutlich, die Parametermenge der Gibbsschen Spezifizierung ist räumlich, nämlich das Gitter, und die Parametermenge der Markovschen Halbgruppe ist in der Regel die Zeit, also eindimensional.

Das Kapitel beende ich mit einem Beispiel für Markov Prozesse: Ich betrachte einen zellularen Automaten (Probabilistic Cellular Automata, PCA). Wir haben wieder eine abzählbare Menge S als Gitter, einen endlichen Zustandsraum, z. B. $E = \{0, 1\}$ oder $E = \{-1, +1\}$, und den Konfigurationsraum $\Omega = E^S$. Die Parametermenge sei abzählbar und stelle die Zeit dar. Ich gebe dann eine spezielle Familie von Wahrscheinlichkeitskernen an, die die Übergangswahrscheinlichkeit von einem Spin $\sigma_{n,i}$ an der Stelle $i \in S$ zur Zeit n zum Spin $\sigma_{n+1,i}$ an der Stelle $i \in S$ zur Zeit $n + 1$ definiert. Diese Markovsche Halbgruppe P definiert einen Markov-Prozeß, der interessante stationäre Maße hat. Ein Maß μ heißt stationäres Maß zum Prozeß, wenn $\mu = \mu P$ gilt. Existenz und Eindeutigkeit eines stationären Maßes sind nicht gesichert. Es wird sich zeigen, dass einige der stationären Maße Gibbs-Maße sind. Es gibt allerdings sowohl stationäre Maße, die keine Gibbs-Maße sind, als auch Gibbs-Maße, die nicht stationär sind.

In Kapitel 4 bearbeite ich Markov Prozesse mit endlichem Zustandsraum E und abzählbarer Parametermenge, so genannte Markov-Ketten. Es wird der Ergodensatz formuliert, der uns Bedingungen für die eindeutige Existenz von stationären Maßen der Markov-Kette liefert.

Wir lernen dann neben dem Gibbs-Feld ein weiteres stochastisches Feld, das Markov-Feld, kennen und werden sehen, dass Markov-Felder und Gibbs-Felder äquivalent sind. Allerdings gilt nicht nur die Äquivalenz von Gibbs- und Markov-Feldern, sondern man kann eine Verbindung zwischen Markov-Feldern und -Ketten zeigen. Wählt man nämlich eine eindimensionale Parametermenge S , betrachtet man z. B. das eindimensionale Gitter $S = \mathbb{Z}$, und einen endlichen Zustandsraum E für das Markov-Feld, so kann gezeigt werden, dass das Markov-Feld äquivalent zu einer Markov-Kette ist. Für diese Markov-Kette wird die Parametermenge nicht wie üblich als Zeit sondern ebenfalls als eindimensionales Gitter interpretiert. Nach dem Ergodensatz hat die Markov-Kette ein eindeutiges stationäres Maß, welches aufgrund der Äquivalenz von Gibbs- und Markov-Feldern ebenfalls ein eindeutiges Gibbs-Maß ist. Als Beispiel führe ich das bekannte eindimensionale Ising-Modell vor, das aufgrund der Eindeutigkeit des Gibbs-Maßes keinen Phasenübergang beschreiben kann. Des Weiteren gebe ich eine Anwendung der Äquivalenz von Markov-Ketten und Gibbs-Maßen aus der Biologie an. Ich betrachte Nukleotidsequenzen der DNS.

Abschließend gehe ich in Kapitel 5 auf Markov-Ketten als Unterrichtsthema der gymnasialen Oberstufe ein.

Teil I.

Gibbs-Maße

2. DLR-Gleichung

Ich beginne mit der Idee von Dobrushin, Lanford und Ruelle, die zur Definition von Gibbs-Maßen mit Hilfe eines bedingten Wahrscheinlichkeitsmaßes führt. Die definierende Gleichung wird nach den Ideengebern DLR-Gleichung genannt, vgl. [12], [13], [14], [44] und [63].

2.1. Grundlagen

Ich werde für die folgenden Erklärungen ein Gittergas als Beispiel heranziehen, siehe [26]. Stellen wir uns also ein Gitter, z. B. \mathbb{Z}^3 , vor. Jeder Gitterpunkt kann im einfachsten Fall von einem Teilchen besetzt oder unbesetzt sein. Ein besetzter Gitterpunkt befindet sich im Zustand 1, ein unbesetzter im Zustand 0. Unser Ziel ist es, ein mathematisches Modell für den Gleichgewichtszustand des Gittergases aufzustellen.

Zunächst benötigen wir einige Vereinbarungen zur Schreibweise. Die vorerst endliche Menge von Gitterpunkten bezeichne ich mit S . Im obigen Beispiel wäre $S \subset \mathbb{Z}^3$. Ein Gitterpunkt kann unterschiedliche Zustände annehmen. Die Menge aller Zustände soll E heißen. Lassen wir nur die Zustände besetzt, also 1, und unbesetzt, 0, zu, so gilt $E = \{0, 1\}$. Alle Zustände, die das gesamte Gitter annehmen kann sind Elemente des Konfigurationsraums $\Omega = E^S$. Ein Element des Raums $(\omega_i)_{i \in S} \in \Omega$ beschreibt also eine Konfiguration des Systems.

Unser System ist sehr komplex und benötigt viele Angaben zur vollständigen Beschreibung, während man für die makroskopische Beschreibung nur wenige Daten verlangen muss. Als Folge des Gesetzes der großen Zahlen kann die Komplexität des Gittergases eingeschränkt werden. Wir definieren ein Maß auf dem Konfigurationsraum.

In das Maß muss z. B. die Wechselwirkung der Teilchen eingebaut werden. Dies berücksichtigen wir mit der Hamilton-Funktion, die z. B. die Form

$$H(\omega) = - \sum_{\{i,j\} \subset S} J(i,j) \omega_i \omega_j - h \sum_{i \in S} \omega_i \quad (2.1)$$

annimmt. Das Interaktionspotential wird mit J und das chemische Potential mit h bezeichnet. Das chemische Potential gibt die Arbeit an, welche benötigt wird, um ein Teilchen in das System zu bringen. Die Hamilton-Funktion H liefert die Gesamtenergie einer Konfiguration.

Nun liefert uns die statistische Mechanik für ein System im Gleichgewicht mit Hamilton-Funktion das endliche Gibbs-Maß

$$\mu(d\omega) = Z^{-1} \exp(-\beta H(\omega)) d\omega, \quad (2.2)$$

2. DLR-Gleichung

wobei β eine positive, zum Kehrwert der absoluten Temperatur proportionale Größe ist. Die Gleichung (2.2) stimmt mit (1.1) überein, wenn wir als a priori Maß auf Ω das Zählmaß wählen, welches sich im Fall $\Omega = \{0, 1\}^S$, $S \subset \mathbb{Z}^3$ ist endlich, anbietet. Für Gittergase beschreibt (2.2) die großkanonische Verteilung, im Fall des Ising-Modells, $E = \{-1, 1\}$, bleibt die Teilchenzahl konstant und wir erhalten die kanonische Verteilung. Die Zustände -1 und $+1$ repräsentieren die Orientierung der Spins.

Dass (2.2) tatsächlich das gesuchte Maß ist, stelle ich an dieser Stelle als Postulat auf. Eine rigorose Rechtfertigung ist Sache der Thermodynamik bzw. statistischen Physik und bis heute nicht vollständig möglich, vielmehr benötigt man dann an anderer Stelle Postulate, vgl. Kapitel 1. Erfolgreiche Anwendungen rechtfertigen allerdings die postulierte Gleichung (2.2).

Unser bisheriges Modell bekommt an einer entscheidenden Stelle ein Problem. Betrachten wir ein System mit einer endlichen Anzahl von Teilchen, so kann das System keinen Phasenübergang beschreiben, da (2.2) analytisch von der Temperatur abhängt. Es können also keine Unstetigkeiten bei Größen wie der Dichte in Abhängigkeit von der Temperatur auftreten. Genau dies passiert aber bei Phasenübergängen. Die Phasenübergänge können mit unserem Modell beschrieben werden, wenn wir den Grenzwert $|S| \rightarrow \infty$ betrachten. Doch mit diesem Grenzprozeß entsteht das nächste Problem, die Definition der Hamilton-Funktion (2.1) ist nicht mehr sinnvoll. Nun kommt die Idee von Dobrushin, Lanford und Ruelle (DLR) ins Spiel: Wir betrachten bedingte Wahrscheinlichkeiten.

Es sei eine endliche Menge S und eine nichtleere Teilmenge Λ von S mit den Konfigurationen $\zeta \in E^\Lambda$ und $\eta \in E^{S \setminus \Lambda}$ auf Λ und $S \setminus \Lambda$ gegeben. Die kombinierte Konfiguration auf S sei $\zeta\eta$. Jetzt bilden wir mit unserem Wahrscheinlichkeitsmaß (2.2) die bedingte Wahrscheinlichkeit, dass ζ in Λ ist unter der Bedingung, dass η in $S \setminus \Lambda$ auftritt

$$\begin{aligned} \mu(\zeta \text{ in } \Lambda | \eta \text{ in } S \setminus \Lambda) &= \frac{\mu(\zeta\eta \text{ in } S)}{\mu(\eta \text{ in } S \setminus \Lambda)} = \frac{\exp(-\beta H(\zeta\eta))}{\sum_{\tilde{\zeta} \in E^\Lambda} \exp(-\beta H(\tilde{\zeta}\eta))} \\ &= Z_\Lambda(\eta)^{-1} \exp(-\beta H_\Lambda(\zeta\eta)), \end{aligned} \quad (2.3)$$

wobei

$$H_\Lambda(\zeta\eta) = - \sum_{\{i,j\} \subset \Lambda} J(i,j) \zeta_i \zeta_j - \sum_{i \in \Lambda} \zeta_i \left(h \sum_{j \in S \setminus \Lambda} J(i,j) \eta_j \right) \quad (2.4)$$

und

$$Z_\Lambda(\eta) = \sum_{\tilde{\zeta} \in E^\Lambda} \exp(-\beta H(\tilde{\zeta}\eta)) \quad (2.5)$$

die Normierungskonstante ist.

Welchen Vorteil hat diese Darstellung? Gehen wir davon aus, dass J eine begrenzte Reichweite hat, so ist (2.4) auch dann noch wohl definiert, wenn S nicht endlich ist. Dies war selbst bei einer begrenzten Reichweite von J für (2.1) nicht gegeben. Die begrenzte Reichweite von J ist eine physikalisch sinnvolle Einschränkung und bedeutet, dass die Konfiguration in Λ nur von dem an Λ grenzenden Teil von $S \setminus \Lambda$ abhängt. Wir können

also über bedingte Wahrscheinlichkeiten den Begriff des Gibbs-Maßes auf unendliche Gitter verallgemeinern. Dies werde ich in den folgenden Abschnitten mathematisch rigoros darstellen.

2.2. Wahrscheinlichkeitskerne und Spezifizierungen

Jetzt sei S eine abzählbar unendliche Menge und (E, \mathfrak{E}) ein Meßraum. $\mathfrak{F} = \mathfrak{E}^S$ sei die Produkt σ -Algebra auf dem Konfigurationsraum $\Omega = E^S = \{\omega = (\omega_i)_{i \in S} : \omega_i \in E\}$. Damit ist \mathfrak{F} die kleinste σ -Algebra auf Ω , die die Zylinderereignisse

$$\{\sigma_\Lambda^{-1}(A)\}, \quad \Lambda \in \mathcal{S}, \quad A \in \mathfrak{E}^\Lambda$$

enthält. Durch σ_i und σ_Λ seien die Projektionen

$$\sigma_i : \Omega \rightarrow E, \quad \omega \mapsto \omega_i$$

und

$$\sigma_\Lambda : \Omega \rightarrow E^\Lambda, \quad \omega \mapsto \omega_\Lambda = (\omega_i)_{i \in \Lambda}$$

definiert. Mit \mathcal{S} bezeichne ich die Menge $\{\Lambda \subset S : 0 < |\Lambda| < \infty\}$ und $\mathfrak{T}_\Lambda = \mathfrak{F}_{S \setminus \Lambda}$ mit $\Lambda \in \mathcal{S}$ ist die σ -Algebra, die die Informationen über die Randbedingungen von Λ enthält. Sie übernimmt die Funktion der σ -Algebra der Vergangenheit eines Markov-Prozesses. Denn hier haben wir keine zeitliche Ordnung. Die σ -Algebra $\mathfrak{F}_{S \setminus \Lambda}$ wird durch die Ereignisse

$$\{\sigma_\Delta^{-1}(A)\}, \quad \Delta \in \mathcal{S}, \quad \Delta \subset S \setminus \Lambda, \quad A \in \mathfrak{E}^\Delta$$

erzeugt. Dass $|\Lambda| < \infty$ gilt, ist wesentlich für mein Ziel, wie ich bereits im letzten Abschnitt mit Blick auf die Hamilton-Funktion angedeutet habe.

Eine \mathfrak{F}_Λ -messbare Funktion wird lokal oder mikroskopisch genannt für $\Lambda \in \mathcal{S}$. Eine Funktion beschreibt eine makroskopische Eigenschaft, wenn sie nicht von Konfigurationen in endlichen Bereichen $\Lambda \in \mathcal{S}$ abhängig ist, wenn sie also \mathfrak{T} -messbar ist mit der terminalen σ -Algebra

$$\mathfrak{T} = \bigcap_{\Lambda \in \mathcal{S}} \mathfrak{T}_\Lambda.$$

Angemerkt sei, dass \mathfrak{T}_Λ und \mathfrak{T} Sub- σ -Algebren der σ -Algebra \mathfrak{F} sind. Wir werden im wesentlichen stochastische Prozesse und stochastische Felder betrachten und beide ausführlich erläutern. Dabei werden wir auch eine Äquivalenz von Markov-Ketten und Gibbs-Feldern kennen lernen. Da wir uns zu Beginn mit der Definition von Gibbs-Maßen beschäftigen, möchte ich die Definition von stochastischen Feldern vorwegnehmen.

Definition 2.2.1. Das Quadrupel $(\Omega, \mathfrak{F}, \mu, (\pi_i)_{i \in S})$ heißt stochastisches Feld, wobei $(\pi_i)_{i \in S}$ eine Familie von Zufallsvariablen auf dem Wahrscheinlichkeitsraum $(\Omega, \mathfrak{F}, \mu)$ mit Werten im gemeinsamen Zustandsraum (E, \mathfrak{E}) ist. S ist die Parametermenge.

Jetzt beginne ich, die für Gibbs-Maße wichtigen Begriffe des Wahrscheinlichkeitskerns und der Spezifizierung zu definieren.

2. DLR-Gleichung

Definition 2.2.2. (X, \mathfrak{X}) und (Y, \mathfrak{Y}) seien Meßräume.

Die Abbildung $K : \mathfrak{X} \times \mathfrak{Y} \rightarrow [0, \infty]$ heißt Kern von \mathfrak{Y} nach \mathfrak{X} , wenn:

- $K(\cdot | y)$ ist Maß auf $(X, \mathfrak{X}) \forall y \in Y$,
- $K(A | \cdot)$ ist \mathfrak{Y} -messbar $\forall A \in \mathfrak{X}$.

Wenn $K(X | \cdot) = 1$, heißt K Wahrscheinlichkeitskern (W-Kern).

Ist \mathfrak{Y} eine Unter- σ -Algebra von \mathfrak{X} , dann heißt der Kern von \mathfrak{Y} nach \mathfrak{X} wohldefiniert, wenn

$$K(A \cap B | \cdot) = K(A | \cdot) 1_B, \quad A \in \mathfrak{X}, B \in \mathfrak{Y}$$

gilt.

Ein Kern K von \mathfrak{Y} nach \mathfrak{X} ordnet jedem Maß μ auf (Y, \mathfrak{Y}) durch

$$\mu K(A) = \int d\mu K(A | \cdot), \quad A \in \mathfrak{X}$$

ein Maß μK auf (X, \mathfrak{X}) zu.

Wahrscheinlichkeitskerne dienen bei dem Markovschen Prozeß zur Definition der Markovschen Halbgruppen, auf die ich später eingehen werde. Man betrachtet durch die Zeit indizierte Familien von Kernen, vgl. [6]. Für die Definition der Spezifizierung gehen wir ebenfalls von einer Familie von Wahrscheinlichkeitskernen aus, indizieren Sie aber mit $\Lambda \in \mathcal{S}$.

Definition 2.2.3. Eine Spezifizierung ist eine Familie $\gamma = (\gamma_\Lambda)_{\Lambda \in \mathcal{S}}$ von wohldefinierten W-Kernen γ_Λ von \mathfrak{T}_Λ nach \mathfrak{F} mit der Parametermenge \mathcal{S} und dem Zustandsraum (E, \mathfrak{E}) , die die Konsistenzbedingung $\gamma_\Delta \gamma_\Lambda = \gamma_\Delta$ für $\Lambda \subset \Delta$ erfüllen, d. h. für \mathfrak{F} -messbare Funktionen und $\omega \in \Omega$ gilt

$$\int \int h(\zeta) \gamma_\Lambda(d\zeta | \sigma) \gamma_\Delta(d\sigma | \omega) = \int h(\sigma) \gamma_\Delta(d\sigma | \omega).$$

Eine Familie von Wahrscheinlichkeitskernen definiert bzw. spezifiziert eine Menge von Wahrscheinlichkeitsmaßen

$$\mathcal{G}(\gamma) = \{\mu \in \mathcal{P}(\Omega, \mathfrak{F}) : \mu(A | \mathfrak{T}_\Lambda) = \gamma_\Lambda(A | \cdot) \text{ } \mu\text{-f.s. } \forall A \in \mathfrak{F} \text{ und } \Lambda \in \mathcal{S}\}. \quad (2.6)$$

$\mathcal{P}(\Omega, \mathfrak{F})$ bezeichnet die Menge der Wahrscheinlichkeitsmaße auf dem Konfigurationsraum Ω . $\mathcal{G}(\gamma)$ ist die Menge der Gibbs-Maße, wenn γ die Gibbssche Spezifizierung ist.

Lemma 2.2.1. Für ein Maß μ gilt

$$\mu(A | \mathfrak{T}_\Lambda) = \gamma_\Lambda(A | \cdot) \text{ } \mu\text{-f.s. } \forall A \in \mathfrak{F} \text{ und } \Lambda \in \mathcal{S} \quad (2.7)$$

genau dann, wenn

$$\mu \gamma_\Lambda = \mu.$$

Beweis. Ein Maß erfüllt (2.7), wenn

$$\mu(A \cap B) = \int_B d\mu \gamma_\Lambda(A | \cdot) \text{ mit } A \in \mathfrak{F} \text{ und } B \in \mathfrak{T}_\Lambda.$$

Da γ_Λ wohldefiniert ist, gleicht die rechte Seite dem Ausdruck $\mu \gamma_\Lambda(A \cap B)$. \square

Der Wahrscheinlichkeitskern kann also als bedingte Wahrscheinlichkeit aufgefasst werden. Die Spezifizierung dient dazu, ein zu den üblichen Problemen der Stochastik inverses Problem zu lösen. Wir haben ein System von bedingten Wahrscheinlichkeiten definiert, ohne ein Wahrscheinlichkeitsmaß zu nutzen, und wollen nun ein Wahrscheinlichkeitsmaß finden, dass γ_Λ als \mathfrak{T}_Λ -bedingte Wahrscheinlichkeit besitzt.

2.3. Gibbssche Spezifizierung

Die folgenden Definitionen sollen dazu dienen, die Brücke zwischen dem ersten und zweiten Abschnitt zu schlagen, siehe [26]. Sie sind für eine saubere mathematische Herleitung mit Hilfe von Familien von Wahrscheinlichkeitskernen nötig. Wir suchen die Gibbssche Spezifizierung, bei der nach den Vorüberlegungen die Hamilton-Funktion eine Rolle spielen muss. Um die Hamilton-Funktion einbauen zu können, werde ich den Begriff der Spezifizierung auf λ -Spezifizierungen ausweiten. Vorher lege ich die folgende Notation fest:

Gegeben sei ein Maß $\lambda \in \mathcal{M}(E, \mathfrak{E})$. Mit $\mathcal{M}(E, \mathfrak{E})$ bezeichne ich die Menge aller σ -endlichen Maße auf (E, \mathfrak{E}) und mit $\lambda = (\lambda_\Lambda)_{\Lambda \in \mathcal{S}}$ die Familie von Kernen λ_Λ von \mathfrak{T}_Λ nach \mathfrak{F} mit

$$\lambda_\Lambda(A|\omega) = \lambda^\Lambda(\zeta \in E^\Lambda : \zeta \omega_{S \setminus \Lambda} \in A) \quad (\omega \in \Omega, \Lambda \in \mathcal{S}). \quad (2.8)$$

Es gilt $\lambda_{\Delta \cup \Lambda} = \lambda_\Delta \lambda_\Lambda$.

Definition 2.3.1. Sei $\lambda \in \mathcal{M}(E, \mathfrak{E})$. Eine λ -Modifikation ist eine Familie $\rho = (\rho_\Lambda)_{\Lambda \in \mathcal{S}}$ von messbaren Funktionen $\rho_\Lambda : \Omega \rightarrow [0, \infty]$, so dass die Familie $\rho\lambda = (\rho_\Lambda \lambda_\Lambda)_{\Lambda \in \mathcal{S}}$ eine Spezifizierung ist. Spezifizierungen γ der Form $\gamma = \rho\lambda$ sind λ -Spezifizierungen.

Für $\gamma = \rho\lambda$ ist die Menge der spezifizierten Maße durch

$$\mathcal{G}(\gamma) = \{\mu \in \mathcal{P}(\Omega, \mathfrak{F}) : \mu = \rho_\Lambda(\mu \lambda_\Lambda), \Lambda \in \mathcal{S}\}.$$

gegeben. Dies folgt aus (2.2.3) mit

$$\rho_\Lambda(\mu \lambda_\Lambda)(A) = \int \mu(d\omega) \int_A \lambda_\Lambda(d\zeta|\omega) \rho_\Lambda(\zeta) = \mu \gamma_\Lambda(A) \quad \forall \mu \in \mathcal{P}(E, \mathfrak{E}), \forall A \in \mathfrak{F}, \forall \Lambda \in \mathcal{S}.$$

Wenn $\tilde{\lambda} = r\lambda$ gilt, wobei r eine messbare Funktion auf E mit $r > 0$ ist, dann ist

$$\tilde{\rho}_\Lambda(\omega) = \frac{\rho_\Lambda(\omega)}{\prod_{i \in \Lambda} r(\omega_i)}, \quad \Lambda \in \mathcal{S}, \omega \in \Omega$$

2. DLR-Gleichung

eine $\tilde{\lambda}$ -Modifikation mit $\tilde{\rho}\tilde{\lambda} = \rho\lambda$. Das bedeutet, dass durch eine geeignete Wahl von r , nämlich mit

$$\int r \, d\lambda = 1,$$

jede λ -Spezifizierung eine $\tilde{\lambda}$ -Spezifizierung für ein $\tilde{\lambda} \in \mathcal{P}(E, \mathfrak{E})$ ist. Die λ -Spezifizierungen $(\rho_\Lambda \lambda_\Lambda)_{\Lambda \in \mathcal{S}}$ sind natürlich normiert, es gilt

$$\rho_\Lambda \lambda_\Lambda = \int \rho_\Lambda(\omega) \lambda_\Lambda(d\omega \mid \cdot) = 1.$$

An dieser Stelle möchte ich ohne Beweis ein Lemma über λ -Modifikationen einfügen, das zunächst etwas unmotiviert zu sein scheint, allerdings bald hilfreich sein wird, siehe [26].

Lemma 2.3.1. *Sei $\lambda \in \mathcal{M}(E, \mathfrak{E})$ und $\rho = (\rho_\Lambda)_{\Lambda \in \mathcal{S}}$ eine Familie von nichtnegativen messbaren Funktionen auf Ω mit $\lambda_\Lambda \rho_\Lambda = 1$. Dann sind die folgenden Aussagen äquivalent:*

- ρ ist eine λ -Modifikation.
- $\forall \Lambda, \Delta \in \mathcal{S}$ mit $\Lambda \subset \Delta$ und $\forall \omega \in \Omega$

$$\rho_\Delta = \rho_\Lambda \lambda_\Lambda \rho_\Delta \quad \lambda_\Delta(\cdot \mid \omega) - f.s..$$

- $\forall \Lambda, \Delta \in \mathcal{S}$ mit $\Lambda \subset \Delta$, $\forall \alpha \in \Omega$ und $\lambda_{\Delta \setminus \Lambda}(\cdot \mid \alpha)$ -fast alle $\omega \in \Omega$ gilt

$$\rho_\Delta(\zeta) \rho_\Lambda(\eta) = \rho_\Delta(\eta) \rho_\Lambda(\zeta)$$

für $\lambda_\Lambda(\cdot \mid \omega) \times \lambda_\Lambda(\cdot \mid \omega)$ -fast alle $(\zeta, \eta) \in \Omega \times \Omega$.

Kommen wir nun zu unserem eigentlichen Ziel, mit der λ -Modifikation die Hamilton-Funktion in die Spezifizierung einzubauen, zurück. Da diese Funktion, wie wir im ersten Abschnitt des Kapitels gesehen haben, wesentlich vom Potential abhängt, werde ich mit folgender Definition dafür sorgen, dass die Hamilton-Funktion wohl definiert ist. Der zu Beginn eingeführte Faktor β soll ab jetzt im Potential Φ enthalten sein.

Definition 2.3.2. Ein Potential ist eine Familie $\Phi = (\Phi_A)_{A \in \mathcal{S}}$ von Funktion $\Phi_A : \Omega \rightarrow \mathbb{R}$ mit den Eigenschaften

- $\forall A \in \mathcal{S}$ gilt Φ_A ist \mathfrak{F}_A -messbar.
- $\forall \Lambda \in \mathcal{S}$ und $\omega \in \Omega$ existiert

$$H_\Lambda^\Phi(\omega) = \sum_{A \in \mathcal{S}, A \cap \Lambda \neq \emptyset} \Phi_A(\omega). \quad (2.9)$$

H_Λ^Φ ist die Hamilton-Funktion in Λ für Φ .

Definition 2.3.3. Ein Potential, dass nur für einelementige und zweielementige Mengen, deren zwei Elemente Nachbarn sind, ungleich 0 ist, heißt Nächster-Nachbar-Potential. Ein Nächster-Nachbar-Potential ist homogen, wenn es zwei Funktionen $\phi_1 : E \times E \mapsto \mathbb{R}$ und $\phi_2 : E \mapsto \mathbb{R}$ gibt, so dass für $\{i, j\} \in B$, $\sigma_i, \sigma_j \in E$

$$\Phi_A = \Phi_A = \begin{cases} \phi_1(\sigma_i, \sigma_j) & \text{wenn } A = \{i, j\}, \\ \phi_2(\sigma_i) & \text{wenn } A = \{i\}, \end{cases}$$

wobei $B = \{b \subset S \mid |b| = 2\}$.

Bemerkung 2.3.1. Wir betrachten $S = \mathbb{Z}^d$, dann gilt $B = \{\{i, j\} \subset \mathbb{Z}^d \mid |i - j| = 1\}$. Allgemeiner liegt ein Paar-Potential vor, wenn $\Phi_A = 0$ gilt, falls $|A| > 2$.

Definition 2.3.4. Ein Potential

- besitzt eine endliche/begrenzte Reichweite, wenn für jedes $i \in S$ ein $\Delta_i \in \mathcal{S}$ existiert, so dass $\Phi_A = 0$ außer $A \subset \Delta_i$ für alle $i \in A$.
- heißt translationsinvariant, wenn für $i \in S$ und $A \in \mathcal{S}$

$$\Phi_A = \Phi_{i+A}$$

gilt, wobei $i + A = \{i + j \mid j \in A\}$.

In Anlehnung an die Gibbs-Maße auf einem beschränkten Gitter können wir für jedes Potential Φ eine Familie $(h_\Lambda^\Phi)_{\Lambda \in \mathcal{S}}$ mit $h_\Lambda^\Phi(\omega) = \exp(-H_\Lambda^\Phi(\omega))$ definieren. Diese Familie bildet noch keinen normierten Kern $h_\Lambda^\Phi \lambda_\Lambda$ und ist daher keine λ -Modifikation. Man spricht von einer Prämodifikation. Mit Blick auf Punkt drei des letzten Lemmas liegt folgende Definition nahe:

Definition 2.3.5. Eine Familie $h = (h_\Lambda)_{\Lambda \in \mathcal{S}}$ von messbaren Funktionen $h_\Lambda : \Omega \rightarrow [0, \infty[$ heißt Prämodifikation, wenn

$$h_\Delta(\zeta)h_\Lambda(\eta) = h_\Lambda(\zeta)h_\Delta(\eta) \quad (2.10)$$

$\forall \Lambda, \Delta \in \mathcal{S}$ mit $\Lambda \subset \Delta$ und $\forall \zeta, \eta \in \Omega$ mit $\zeta_{S \setminus \Lambda} = \eta_{S \setminus \Lambda}$ gilt.

Lemma 2.3.2. Die Familie $(h_\Lambda^\Phi)_{\Lambda \in \mathcal{S}}$ ist für jedes Potential eine Prämodifikation.

Beweis. Sei $\emptyset \neq \Lambda \subset \Delta \in \mathcal{S}$ und $\zeta, \eta \in \Omega$ mit $\zeta_{S \setminus \Lambda} = \eta_{S \setminus \Lambda}$. Laut Definition einer Prämodifikation muss

$$h_\Delta^\Phi(\zeta)h_\Lambda^\Phi(\eta) = h_\Lambda^\Phi(\zeta)h_\Delta^\Phi(\eta)$$

gezeigt werden. Einsetzen ergibt

$$\exp(-(H_\Delta^\Phi(\zeta) + H_\Lambda^\Phi(\eta))) = \exp(-(H_\Lambda^\Phi(\zeta) + H_\Delta^\Phi(\eta)))$$

und damit

$$\begin{aligned} H_\Delta^\Phi(\zeta) + H_\Lambda^\Phi(\eta) &= H_\Lambda^\Phi(\zeta) + H_\Delta^\Phi(\eta) \\ H_\Delta^\Phi(\zeta) - H_\Lambda^\Phi(\zeta) &= H_\Delta^\Phi(\eta) - H_\Lambda^\Phi(\eta). \end{aligned} \quad (2.11)$$

2. DLR-Gleichung

Für $H_\Delta^\Phi - H_\Lambda^\Phi$ ergibt sich mit (2.3.2) Punkt 2

$$H_\Delta^\Phi - H_\Lambda^\Phi = \sum_{A \cap \Delta \neq \emptyset, A \cap \Lambda = \emptyset} \Phi_A.$$

Die Reihe existiert ebenfalls nach **Definition 2.3.2** Punkt 2. Da des Weiteren $\zeta, \eta \in \Omega$ mit $\zeta_{S \setminus \Lambda} = \eta_{S \setminus \Lambda}$ gilt, ist

$$\sum_{A \cap \Delta \neq \emptyset, A \cap \Lambda = \emptyset} \Phi_A(\zeta) = \sum_{A \cap \Delta \neq \emptyset, A \cap \Lambda = \emptyset} \Phi_A(\eta)$$

und damit (2.11) erfüllt. \square

Um nun die Normierung des Kerns zu erreichen, muss

$$Z_\Lambda^\Phi(\omega) = \lambda_\Lambda h_\Lambda^\Phi(\omega) = \int \lambda^\Lambda(d\zeta) \exp(-H_\Lambda^\Phi(\zeta \omega_{S \setminus \Lambda})) \quad \forall \Lambda \in \mathcal{S} \text{ und } \omega \in \Omega \quad (2.12)$$

endlich sein. Ist die Endlichkeit von (2.12) erfüllt, so ist das Potential Φ λ -zulässig. Mit einem λ -zulässigen Potential kann eine λ -Modifikation konstruiert werden.

Lemma 2.3.3. *Es sei h eine Prämodifikation und für $\lambda \in \mathcal{M}(E, \mathfrak{E})$ gelte $0 < \lambda_\Lambda h_\Lambda < \infty$ für alle $\Lambda \in \mathcal{S}$. Dann ist $\rho = \left(\frac{h_\Lambda}{\lambda_\Lambda h_\Lambda} \right)_{\Lambda \in \mathcal{S}}$ eine λ -Modifikation.*

Beweis. Da $\lambda_\Lambda h_\Lambda$ \mathfrak{T}_Λ -messbar ist, folgt

$$\lambda_\Lambda \rho_\Lambda = \lambda_\Lambda \left(\frac{h_\Lambda}{\lambda_\Lambda h_\Lambda} \right) = \frac{\lambda_\Lambda h_\Lambda}{\lambda_\Lambda h_\Lambda} = 1.$$

Da ρ eine Prämodifikation ist, erfüllt ρ Punkt 3 **Lemma 2.3.1**:

$$\begin{aligned} & \rho_\Delta(\zeta) \rho_\Lambda(\eta) \\ &= \frac{h_\Delta(\zeta) h_\Lambda(\eta)}{\lambda_\Delta h_\Delta(\zeta) \lambda_\Lambda h_\Lambda(\eta)} \stackrel{h_\Delta(\zeta) h_\Lambda(\eta) = h_\Lambda(\zeta) h_\Delta(\eta)}{=} \frac{h_\Delta(\eta) h_\Lambda(\zeta)}{\lambda_\Delta h_\Delta(\zeta) \lambda_\Lambda h_\Lambda(\eta)} \stackrel{\zeta_{S \setminus \Lambda} = \eta_{S \setminus \Lambda}}{=} \frac{h_\Delta(\eta) h_\Lambda(\zeta)}{\lambda_\Delta h_\Delta(\eta) \lambda_\Lambda h_\Lambda(\zeta)} \\ &= \rho_\Delta(\eta) \rho_\Lambda(\zeta) \end{aligned}$$

Zusammen mit Beweisteil 1 folgt die Behauptung. \square

Wir erhalten für ein λ -zulässiges Potential Φ die Funktionen $\rho_\Lambda^\Phi = \frac{h_\Lambda^\Phi}{Z_\Lambda^\Phi}$, $\Lambda \in \mathcal{S}$ auf Ω .

Durch $\rho^\Phi = (\rho_\Lambda^\Phi)_{\Lambda \in \mathcal{S}}$ ist eine λ -Modifikation definiert.

Nun können wir alles zur entscheidenden Definition von Gibbs-Maßen zusammensetzen.

Definition 2.3.6. Sei Φ ein λ -zulässiges Potential und $\omega \in \Omega$, $\Lambda \in \mathcal{S}$. Dann wird das bedingte Wahrscheinlichkeitsmaß

$$A \rightarrow \gamma_\Lambda^\Phi(A|\omega) = \rho_\Lambda^\Phi \lambda_\Lambda(A|\omega) = Z_\Lambda^\Phi(\omega)^{-1} \int \lambda^\Lambda(d\zeta) \exp(-H_\Lambda^\Phi(\zeta \omega_{S \setminus \Lambda})) 1_A(\zeta \omega_{S \setminus \Lambda}) \quad (2.13)$$

auf (Ω, \mathfrak{F}) Gibbsverteilung in Λ mit der Randbedingung $\omega_{S \setminus \Lambda}$ oder DLR-Gleichung genannt. Die λ -Spezifizierung $\gamma^\Phi = (\gamma_\Lambda^\Phi)_{\Lambda \in \mathcal{S}} = \rho^\Phi \lambda$ heisst Gibbssche Spezifizierung für Φ und λ . Jedes Wahrscheinlichkeitsmaß $\mu \in \mathcal{G}(\Phi) = \mathcal{G}(\gamma^\Phi)$ heisst Gibbs-Maß.

Bemerkung 2.3.2. Es gilt:

- Die Gibbs-Maße μ müssen

$$\mu(A) = \int_{\Omega} \mu(\omega) \gamma_{\Lambda}(A | \omega) \quad (2.14)$$

erfüllen. Da μ nur auf $E^{S \setminus \Lambda}$ wirkt, können wir auch

$$\mu(A) = \int_{E^{S \setminus \Lambda}} \mu_{S \setminus \Lambda}(\omega) \gamma_{\Lambda}(A | \omega)$$

schreiben.

- Die Gibbs-Maße liefern ein mathematisches Modell eines Gleichgewichtszustandes. Dies habe ich in der Einleitung dieser Arbeit und in Abschnitt 2.1 motiviert. Wie im Fall der endlichen Gibbs-Maße können auch die Gibbs-Maße mit dem Variationsprinzip oder physikalisch als Maße die die freie Energie minimieren charakterisiert werden, siehe [26].
- Mit Hilfe von Projektionen

$$\pi_{\Lambda} : \Omega \rightarrow E^{\Lambda}$$

können wir auch $\rho_{\Lambda}^{\Phi} \lambda_{\Lambda}(\pi_{\Lambda}^{-1}(\zeta) | \pi_{S \setminus \Lambda}^{-1}(\omega_{S \setminus \Lambda}))$ für die Wahrscheinlichkeit von ζ in Λ unter der Bedingung $\omega_{S \setminus \Lambda}$ in $S \setminus \Lambda$ schreiben. Dabei umfaßt $\pi_{\Lambda}^{-1}(\zeta)$ alle Konfigurationen $\sigma \in \Omega$, deren Einschränkung auf Λ gleich ζ ist:

$$\rho_{\Lambda}^{\Phi} \lambda_{\Lambda}(\pi_{\Lambda}^{-1}(\zeta) | \pi_{S \setminus \Lambda}^{-1}(\omega_{S \setminus \Lambda})) = Z_{\Lambda}^{\Phi}(\omega)^{-1} \exp(-H_{\Lambda}^{\Phi}(\zeta \omega_{S \setminus \Lambda})). \quad (2.15)$$

- Ist das Potential Φ von endlicher Reichweite, so werden endlich viele Gitterpunkte und ein diskreter Wahrscheinlichkeitsraum betrachtet. Das Gibbs-Maß muss nach der Definition der bedingten Wahrscheinlichkeiten $P(A \cap B) = P(A|B)P(B)$ die Bedingung

$$\mu(\pi_{\Lambda}^{-1}(\zeta) \cap \pi_{S \setminus \Lambda}^{-1}(\omega_{S \setminus \Lambda})) = \rho_{\Lambda}^{\Phi} \lambda_{\Lambda}(\pi_{\Lambda}^{-1}(\zeta) | \pi_{S \setminus \Lambda}^{-1}(\omega_{S \setminus \Lambda})) \mu(\pi_{S \setminus \Lambda}^{-1}(\omega_{S \setminus \Lambda}))$$

erfüllen.

Die folgende Bemerkung dient an dieser Stelle als Ausblick und wird noch vertieft.

Bemerkung 2.3.3. Es gilt:

- Ist μ ein Gibbs-Maß, dann ist $(\Omega, \mathfrak{F}, \mu, (\pi_i)_{i \in S})$ ein Gibbs-Feld.
- Ein stochastisches Feld, dessen W-Maß μ die Bedingung

$$\mu(A | \mathfrak{T}_{\Lambda}) = \mu(A | \mathfrak{F}_{\partial \Lambda}) \quad \mu\text{-fast sicher}$$

mit $A \in \mathfrak{F}_{\Lambda}$ und $\Lambda \in \mathcal{S}$ erfüllt, heißt Markov-Feld.

- Markov-Felder und Gibbs-Felder sind äquivalent.

2.4. Existenz und Phasenübergänge

Die Menge der Gibbs-Maße $\mathcal{G}(\gamma^\Phi)$ kann die leere Menge sein, sie kann aber auch mehr als nur ein Element beinhalten. Darauf will ich in diesem Abschnitt eingehen. Vorweg möchte ich eine physikalische Interpretation liefern. Die Gibbs-Maße sind Wahrscheinlichkeitsmaße auf dem Konfigurationsraum Ω . Sie sollen den Gleichgewichtszustand eines Gittergases beschreiben. Enthält $\mathcal{G}(\gamma^\Phi)$ mehr als ein Element, so kann das System nach unserem Modell zwischen verschiedenen Gleichgewichtszuständen wählen. Diese Wahlfreiheit charakterisiert einen Phasenübergang.

Definition 2.4.1. Ein Potential lässt einen Phasenübergang zu, wenn $|\mathcal{G}(\gamma^\Phi)| > 1$.

Im Folgenden sei Φ ein translationsinvariantes Potential mit begrenzter Reichweite. Das Ziel dieses Abschnitts ist der Beweis des folgenden Satzes, vgl. [40].

Satz 2.4.1. $\mathcal{G}(\gamma^\Phi)$ ist eine nichtleere, kompakte und konvexe Menge.

Bemerkung 2.4.1. Tatsächlich kann $\mathcal{G}(\gamma^\Phi)$ auch die leere Menge sein. Dies ist beispielsweise für das diskrete Gaußsche Modell im eindimensionalen und für das Gaußsche Modell im ein- und zweidimensionalen der Fall. In diesen Modellen ist der Zustandsraum $E = \mathbb{Z}$ bzw. $E = \mathbb{R}$ und damit nicht mehr endlich. Ich gehe im folgenden Beweis von einem endlichen Zustandsraum aus, z. B. $E = \{0, 1\}$ für Gittergase oder $E = \{-1, +1\}$ für das Isingmodell.

Da nun einige topologische Ausführungen kommen, benötige ich zwei Definitionen.

Definition 2.4.2. Ein topologischer Raum ist ein Paar (X, \mathcal{O}) , bestehend aus einer Menge X und einer Familie \mathcal{O} von Teilmengen (genannt „offene Mengen“) von X , derart dass gilt:

- Beliebige Vereinigungen von offenen Mengen sind offen.
- Der Durchschnitt von je zwei offenen Mengen ist offen.
- \emptyset und X sind offen.

\mathcal{O} ist die Topologie von X .

Definition 2.4.3. Eine Topologie \mathcal{O}_1 auf X heißt gröber als eine Topologie \mathcal{O}_2 auf X , wenn $\mathcal{O}_1 \subset \mathcal{O}_2$ gilt, also \mathcal{O}_1 weniger offene Mengen aufweist als \mathcal{O}_2 .

Das Beweisprinzip für die Existenz von Gibbs-Maßen ist die Wahl einer Topologie, so dass jede Folge von Wahrscheinlichkeitsmaßen auf dem Konfigurationsraum Ω einen Häufungspunkt besitzt, der ein Gibbs-Maß ist. Dabei sind die Folgenglieder keine Gibbs-Maße. Dies stellt an die Topologie von $\mathcal{P}(\Omega)$ die Ansprüche

- so grob zu sein, dass $\mathcal{P}(\Omega, \mathfrak{F})$ kompakt ist und
- so fein zu sein, dass die Abbildungen $\mu \mapsto \int_\Omega f d\mu$ für jede beschränkte messbare Funktion f stetig sind.

Diese entgegengesetzten Forderungen erfüllt die Topologie der lokalen Konvergenz. Ist E ein begrenzter Zustandsraum, so fällt die Topologie der lokalen Konvergenz mit der schwach- $*$ -Topologie zusammen. In der **Bemerkung 2.4.1** habe ich mich bereits auf endliche Zustandsräume festgelegt, um $\mathcal{G}(\gamma^\Phi) = \emptyset$ auszuschließen. Ich werde also die schwach- $*$ -Topologie nutzen.

Definition 2.4.4. Die schwach- $*$ -Topologie auf $\mathcal{P}(\Omega)$ ist die grösste Topologie auf $\mathcal{P}(\Omega)$, für die die Abbildungen $\mu \mapsto \int_\Omega f d\mu$ für alle $f \in C(\Omega)$ stetig sind.

Auf E wähle ich die diskrete Topologie, d. h. die Topologie, in der alle Teilmengen von E offen sind. Die diskrete ist die feinste Topologie. Auf $\Omega = E^S$ nutzen wir die Produkttopologie.

Definition 2.4.5. Die Produkttopologie ist die grösste, unter der die Projektionen

$$\pi_l : \Omega \rightarrow E, \quad \omega \mapsto \omega_l, \quad l \in S$$

stetige Abbildungen sind. Eine Abbildung ist stetig, wenn Urbilder offener Mengen stets offen sind.

Bei uns ist E mit der diskreten Topologie ein kompakter Raum. Nach dem Satz von Tychonoff sind beliebige Produkte kompakter topologischer Räume in der Produkttopologie kompakt. Also ist Ω mit der Produkttopologie kompakt. Dies wäre mit der diskreten Topologie auf Ω nicht der Fall. Einelementige Teilmengen von Ω sind im Gegensatz zur diskreten Topologie mit der Produkttopologie nicht offen.

Bemerkung 2.4.2. In der **Bemerkung 2.4.1** habe ich darauf hingewiesen, dass $\mathcal{G}(\gamma^\Phi)$ leer sein kann. Für z. B. $E = \mathbb{Z}$ ist E mit der diskreten Topologie ausgestattet nicht mehr kompakt. Damit gilt der Satz von Tychonoff nicht mehr und Ω ist nicht kompakt. Der folgende Beweis für **Satz 2.4.1** funktionierte nicht!

Ich knüpfe nun an die beiden Forderungen an die Topologie auf $\mathcal{P}(\Omega)$ an. Die zweite haben wir mit der Wahl der schwach- $*$ -Topologie erfüllt. Es bleibt die erste Forderung.

Satz 2.4.2. In der schwach- $*$ -Topologie ist $\mathcal{P}(\Omega)$ kompakt.

Beweis. Ω ist nach den obigen Ausführungen kompakt. Des Weiteren ist Ω ein Hausdorffraum, d. h. alle $\omega^1 \neq \omega^2 \in \Omega$ besitzen zwei disjunkte Umgebungen, da für zwei Konfigurationen $\omega^1 \neq \omega^2 \in \Omega$ ein Gitterpunkt $k \in S$ mit $\omega_k^1 \neq \omega_k^2 \in \Omega$ existiert. Damit sind die $\pi_k^{-1}(\omega_k^i) \subset \Omega$ disjunkte Umgebungen von ω^i .

Der Satz von Riesz-Markov sagt, dass sich für einen kompakten Hausdorffraum Ω alle positiven linearen Funktionale $l : C(\Omega) \rightarrow \mathbb{C}$ mit einem eindeutigen Bairemaß μ auf Ω mit $l(f) = \int_\Omega f d\mu$ identifizieren lassen, vgl. [60]. $C(\Omega)$ ist der Raum der stetigen Funktionen auf Ω .

Wir können nach dem Satz von Riesz-Markov die positiven, linearen Identitätserhaltenden Funktionale mit den Wahrscheinlichkeitsmaßen identifizieren. Damit können die Wahrscheinlichkeitsmaße als Elemente des Dualraums von $C(\Omega)$ betrachtet werden. Die

2. DLR-Gleichung

Menge aller stetigen linearen Funktionale von $C(\Omega)$ nach \mathbb{C} ist der Dualraum $(C(\Omega))^*$ von Ω . Die Wahrscheinlichkeitsmaße bilden eine abgeschlossene (Eine Folge von Maßen, deren Wert für den gesamten Raum 1 ist, hat auch solch ein Maß als Grenzwert) konvexe Menge in der Einheitskugel (da normiert!) von $(\Omega(\Omega))^*$. Die Einheitskugel des Dualraums eines Banachraumes ist nach dem Satz von Banach-Alaoglu schwach- $*$ -kompakt, siehe [60]. Da $C(\Omega)$ ein Banachraum ist (stetige Funktionen über Kompaktum mit Supremumsnorm), ist die Einheitskugel von $(C(\Omega))^*$ schwach- $*$ -kompakt. Aufgrund der Identifizierung der Funktionale mit den Wahrscheinlichkeitsmaßen und der Abgeschlossenheit in $(C(\Omega))^*$ folgt, dass $\mathcal{P}(\Omega)$ ebenfalls schwach- $*$ -kompakt ist. \square

Bemerkung 2.4.3. Unsere Observablen sind stetige Funktionen $O : \Omega \mapsto \mathbb{R}$. Mit den Wahrscheinlichkeitsmaßen auf Ω lassen sich die Erwartungswerte $\langle O \rangle = \int_{\Omega} O(\omega) d\mu(\omega)$ mit $\mu \in \mathcal{P}(\Omega)$ definieren. Das Funktional $O \rightarrow \langle O \rangle$ vom Vektorraum der stetigen Funktionen auf Ω in den Körper der reellen Zahlen ist linear, positiv für nichtnegative stetige Funktionen und identitätserhaltend. Sie bilden eine konvexe Menge, da für zwei lineare, positive und identitätserhaltende Funktionale μ_1, μ_2 und für $0 \leq \lambda \leq 1$

$$\lambda\mu_1 + (1 - \lambda)\mu_2$$

ebenfalls ein lineares, positives und identitätserhaltendes Funktional auf dem Raum der Observablen ist.

Ich werde mit diesen topologischen Grundlagen jetzt zeigen, dass $\mathcal{G}(\gamma^\Phi)$ nicht leer ist. Mit der folgenden Vorgehensweise lassen sich alle Gibbs-Maße konstruieren. Ich beginne mit dem ersten Teil des Beweises von **Satz 2.4.1**.

Ich wähle eine Folge $(\Lambda_i)_{i \in \mathbb{N}}$ von endlichen Teilmengen $\Lambda_i \in \mathcal{S}$, die gegen die ganze Menge S konvergiert:

- $\Lambda_i \subset \Lambda_{i+1}$ für $i \in \mathbb{N}$.
- Für alle $\Lambda \in \mathcal{S}$ existiert ein Λ_i mit $\Lambda \subset \Lambda_i$.

In Anlehnung an die Definition von Gibbs-Maßen können wir den Konfigurationsraum in E^{Λ_i} und $E^{S \setminus \Lambda_i}$ aufspalten. Jetzt definiere ich eine Folge $(P_i)_{i \in \mathbb{N}}$ mit $P_i \in \mathcal{P}(\Omega)$

$$P_i(\{\omega\}) = \frac{\exp(-H_{\Lambda_i}^\Phi(\zeta\omega_{S \setminus \Lambda_i}))}{\int \lambda^{\Lambda_i}(d\zeta) \exp(-H_{\Lambda_i}^\Phi(\zeta\omega_{S \setminus \Lambda_i}))}, \quad (2.16)$$

wobei $\omega = (\zeta\omega_{S \setminus \Lambda_i})$ die Konfiguration auf S mit einer Randbedingung $\omega_{S \setminus \Lambda_i} \in E^{S \setminus \Lambda_i}$ sei.

Da $\mathcal{P}(\Omega)$ laut **Satz 2.4.2** schwach- $*$ -kompakt ist, hat die Folge $(P_i)_{i \in \mathbb{N}}$ für $i \rightarrow \infty$ einen Häufungspunkt P , gegen den eine Teilfolge konvergiert. Mit dem nächsten Satz erhalten wir $\mathcal{G}(\gamma^\Phi) \neq \emptyset$.

Satz 2.4.3. P ist bezüglich des Potentials Φ ein Gibbs-Maß.

Beweis. Zu Beginn dieses Abschnitts habe ich Potentiale mit begrenzter Reichweite vorausgesetzt. Sei die Reichweite durch R gegeben. Betrachten wir ein Gebiet $\Lambda \in \mathcal{S}$ dann reicht es, Randbedingungen für $\Theta := \Gamma - \Lambda$ mit $\Gamma = \{k \in S \mid d(k, \Lambda) \leq R\}$ festzulegen. Ist P ein Gibbs-Maß müssen nach **Bemerkung 2.3.2** für $\theta \in \Theta$ die zum Maß P gehörigen bedingten Wahrscheinlichkeiten die Form

$$P(\pi_\Lambda^{-1}(\zeta) \mid \pi_\Theta^{-1}(\theta)) = \frac{\exp(-H_\Lambda^\Phi(\zeta\theta))}{\int \lambda^\Lambda(d\zeta) \exp(-H_\Lambda^\Phi(\zeta\theta))} \quad (2.17)$$

annehmen.

Für jede Wahl von Λ und damit Γ finden wir ein i_0 mit $\Gamma \subset \Lambda_i \forall i \geq i_0$ (siehe Definition der Folge $(\Lambda_i)_{i \in \mathbb{N}}$).

$\forall i \geq i_0$ gilt

$$P_i(\pi_\Lambda^{-1}(\zeta) \mid \pi_\Theta^{-1}(\theta)) = \frac{P_i(\pi_\Gamma^{-1}(\zeta\theta))}{P_i(\pi_\Theta^{-1}(\theta))}. \quad (2.18)$$

Sehen wir uns den Zähler und Nenner getrennt an:

$$\begin{aligned} P_i(\pi_\Gamma^{-1}(\zeta\theta)) &= \frac{\int \lambda^{\Lambda_i \setminus \Gamma}(d\chi) \exp(-H_{\Lambda_i}^\Phi(\zeta\theta\chi\omega_{S \setminus \Lambda_i}))}{\int \lambda^\Lambda(d\zeta) \int \lambda^\Theta(d\theta') \int \lambda^{\Lambda_i \setminus \Gamma}(d\chi) \exp(-H_{\Lambda_i}^\Phi(\zeta\theta'\chi\omega_{S \setminus \Lambda_i}))} \\ P_i(\pi_\Theta^{-1}(\theta)) &= \frac{\int \lambda^\Lambda(d\zeta) \int \lambda^{\Lambda_i \setminus \Gamma}(d\chi) \exp(-H_{\Lambda_i}^\Phi(\zeta\theta\chi\omega_{S \setminus \Lambda_i}))}{\int \lambda^\Lambda(d\zeta) \int \lambda^\Theta(d\theta') \int \lambda^{\Lambda_i \setminus \Gamma}(d\chi) \exp(-H_{\Lambda_i}^\Phi(\zeta\theta'\chi\omega_{S \setminus \Lambda_i}))}. \end{aligned}$$

Damit können wir den Quotienten (2.18) neu aufschreiben:

$$\begin{aligned} P_i(\pi_\Lambda^{-1}(\zeta) \mid \pi_\Theta^{-1}(\theta)) &= \frac{\int \lambda^{\Lambda_i \setminus \Gamma}(d\chi) \exp(-H_{\Lambda_i}^\Phi(\zeta\theta\chi\omega_{S \setminus \Lambda_i}))}{\int \lambda^\Lambda(d\zeta) \int \lambda^{\Lambda_i \setminus \Gamma}(d\chi) \exp(-H_{\Lambda_i}^\Phi(\zeta\theta\chi\omega_{S \setminus \Lambda_i}))} \\ &= \frac{\int \lambda^{\Lambda_i \setminus \Gamma}(d\chi) \exp(-\sum_{A \in \mathcal{S}, A \cap \Lambda_i \neq \emptyset} \Phi_A(\zeta\theta\chi\omega_{S \setminus \Lambda_i}))}{\int \lambda^\Lambda(d\zeta) \int \lambda^{\Lambda_i \setminus \Gamma}(d\chi) \exp(-\sum_{A \in \mathcal{S}, A \cap \Lambda_i \neq \emptyset} \Phi_A(\zeta\theta\chi\omega_{S \setminus \Lambda_i}))}. \end{aligned}$$

Die Summanden mit $A \cap \Lambda_i \neq \emptyset$ aber $A \cap \Lambda = \emptyset$ tauchen sowohl im Zähler als auch im Nenner im Argument der Exponentialfunktion auf und lassen sich kürzen:

$$\begin{aligned} &= \frac{\exp(-\sum_{A \in \mathcal{S}, A \cap \Lambda \neq \emptyset} \Phi_A(\zeta\theta))}{\int \lambda^\Lambda(d\zeta) \exp(-\sum_{A \in \mathcal{S}, A \cap \Lambda \neq \emptyset} \Phi_A(\zeta\theta))} \\ &= \frac{\exp(-H_\Lambda^\Phi(\zeta\theta))}{\int \lambda^\Lambda(d\zeta) \exp(-H_\Lambda^\Phi(\zeta\theta))}. \end{aligned}$$

Der Quotient (2.18) ist also von i unabhängig und hat die gewünschte Form (2.17). Damit existiert der Grenzwert P und ist ein Gibbs-Maß. \square

Es bleibt für den **Satz 2.4.1** zu zeigen, dass $\mathcal{G}(\gamma^\Phi)$ eine kompakte und konvexe Menge ist.

2. DLR-Gleichung

Beweis. Es ist $\mathcal{G}(\gamma^\Phi)$ eine abgeschlossene Teilmenge der kompakten Menge $\mathcal{P}(\Omega)$, da Häufungspunkte von Gibbs-Maßen wieder Gibbs-Maße sind, und damit kompakt.

Warum sind Häufungspunkte von Gibbs-Maßen wieder Gibbs-Maße?

Betrachten wir diesmal eine Folge von Gibbs-Maßen P_i in der Menge $\mathcal{G}(\gamma^\Phi)$. Es gilt für alle i

$$\frac{P_i(\pi_\Lambda^{-1}(\zeta) \cap \pi_\Theta^{-1}(\theta))}{P_i(\pi_\Theta^{-1}(\theta))} = \frac{\exp(-H_\Lambda^\Phi(\zeta\theta))}{\int \lambda^\Lambda(d\zeta) \exp(-H_\Lambda^\Phi(\zeta\theta))}, \quad (2.19)$$

da alle $P_i \in \mathcal{G}(\gamma^\Phi)$ (siehe **Bemerkung 2.3.2**). Also gilt (2.19) auch für den Häufungspunkt von P_i .

P_0 und P_1 seien Elemente von $\mathcal{G}(\gamma^\Phi)$. Zu zeigen ist, dass dann auch

$$P_\delta := (1 - \delta)P_0 + \delta P_1 \in \mathcal{G}(\gamma^\Phi) \quad \forall \delta \in [0, 1]. \quad (2.20)$$

Da $\mathcal{P}(\Omega)$ eine konvexe Menge ist, ist $P_\delta \in \mathcal{P}(\Omega)$. Die Reichweite des Potentials sei R . Wieder müssen wir nur für $\Theta := \{k \in S \setminus \Lambda \mid d(k, \Lambda) \leq R\}$ mit $\Lambda \in \mathcal{S}$ Randbedingungen festlegen. Es sei $\theta \in E^\Theta$ Randbedingung. Wählen wir $P_0(\pi_\Theta^{-1}(\theta)) > 0$ und $P_1(\pi_\Theta^{-1}(\theta)) > 0$, so ist auch $P_\delta(\pi_\Theta^{-1}(\theta)) > 0$. Dies ermöglicht folgende Umformung. Seien $\zeta \in E^\Lambda$, $A := \pi_\Lambda^{-1}(\zeta)$ und $B := \pi_\Theta^{-1}(\theta)$, dann gilt

$$P_\delta(A|B) = \frac{P_\delta(A \cap B)}{P_\delta(B)} \stackrel{(2.20)}{=} \frac{(1 - \delta)P_0(A \cap B) + \delta P_1(A \cap B)}{(1 - \delta)P_0(B) + \delta P_1(B)}. \quad (2.21)$$

Da P_0 und P_1 Gibbs-Maße sind folgt für $i \in \{0, 1\}$

$$P_i(A \cap B) = \frac{\exp(-H_\Lambda^\Phi(\zeta\theta))}{\int \lambda^\Lambda(d\zeta) \exp(-H_\Lambda^\Phi(\zeta\theta))} \cdot P_i(B). \quad (2.22)$$

Setzt man (2.22) in (2.21) ein, erhält man

$$P_\delta(\pi_\Lambda^{-1}(\zeta) | \pi_\Theta^{-1}(\theta)) = \frac{\exp(-H_\Lambda^\Phi(\zeta\theta))}{\int \lambda^\Lambda(d\zeta) \exp(-H_\Lambda^\Phi(\zeta\theta))}.$$

Damit ist $P_\delta \in \mathcal{G}(\gamma^\Phi)$ und der **Satz 2.4.1** bewiesen. □

Definition 2.4.6. Ein Gibbs-Maß P heißt extremal, wenn es keine Zustände $P_1 \neq P_2$ mit

$$P = (1 - \delta)P_1 + \delta P_2 \quad \text{und} \quad \delta \in [0, 1]$$

gibt.

Bemerkung 2.4.4. Der Satz von Krein-Milman besagt, dass eine nicht leere kompakte konvexe Teilmenge K eines lokal-konvexen Raumes die abgeschlossene konvexe Hülle ihrer Extremalpunkte $Ex(K)$ ist. Die Menge $\mathcal{G}(\gamma^\Phi)$ lässt sich also aus ihren Extremalpunkten rekonstruieren. Die extremalen Punkte der Menge $\mathcal{G}(\gamma^\Phi)$ bezeichnet man als reine Gibbs-Maße.

Diese reinen Gibbs-Maße sind von besonderem Interesse. In Experimenten zeigt sich nämlich, dass die makroskopischen Maßgrößen schwankungsfrei sind. Nutzt man reine Gibbs-Maße, verschwindet die Varianz der makroskopischen Größen.

Die reinen Gibbs-Maße werden als Phasen bezeichnet. So existieren für Wasser am Gefrierpunkt die Phasen Wasser und Eis. Auf Wasser schwimmendes Eis wird durch ein weiteres reines, nicht translationsinvariantes Gibbs-Maß, beschrieben, aber nicht durch eine Kombination der reinen Maße für die Phasen Wasser und Eis. Eine Kombination beschreibt die Unkenntnis über den wirklichen Zustand!

Zu Beginn des Abschnitts habe ich definiert, dass das Potential Φ für $|\mathcal{G}(\gamma^\Phi)| > 1$ einen Phasenübergang zulässt. Es stellt sich nun nach dem Existenzbeweis die Frage, ob und wann wir eine Abwesenheit von Phasenübergängen haben, das Gibbs-Maß also eindeutig ist. Hierzu verweise ich auf [26].

Wäre das Potential gleich Null, hätten wir nur das Produktmaß. Man kann nun zeigen, dass für betragsmäßig kleine Potentiale nur ein Gibbs-Maß existiert. Für hohe Temperaturen gibt es folglich auch nur ein Gibbs-Maß, da $\beta = \frac{1}{kT}$ als Faktor ins Potential gezogen wurde.

Der später in Kapitel 4 ausführlich diskutierte eindimensionale Fall liefert ebenfalls die Abwesenheit eines Phasenübergangs. Ich werde für diesen Fall den Zusammenhang zu Markov-Ketten herausarbeiten. Dennoch war das eindimensionale Modell, auch Ising Modell genannt, der erste Schritt zur mathematischen Theorie der Phasenübergänge, da in höheren Dimensionen das Gibbs-Maß nicht länger eindeutig ist.

2. *DLR-Gleichung*

3. Markov-Prozeß und zellulare Automaten

In diesem Kapitel möchte ich an erster Stelle stochastische Prozesse definieren, um später die Verbindung zu stochastischen Feldern, insbesondere zwischen Markov-Ketten und Markov- bzw. Gibbs-Feldern, herstellen zu können.

Wir betrachten dann bedingte Gibbs-Maße als stationäre Maße einer kontinuierlichen Markov-Kette und ich führe zellulare Automaten ein, die ein Beispiel für Markov-Prozesse liefern.

3.1. Stochastischer Prozeß und Markovsche Halbgruppe

Definition 3.1.1. Stochastischer Prozeß heißt jedes Quadrupel $(\dot{\Omega}, \dot{\mathfrak{F}}, \dot{P}, (X_t)_{t \in T})$, wobei $(\dot{\Omega}, \dot{\mathfrak{F}}, \dot{P})$ ein Wahrscheinlichkeitsraum und $(X_t)_{t \in T}$ eine Familie von Zufallsvariablen auf diesem W-Raum mit Werten in einem gemeinsamen Meßraum (Ω, \mathfrak{F}) ist. $T \subset \mathbb{R}$ ist die Parameter- oder Zeitmenge und (Ω, \mathfrak{F}) der Zustandsraum des Prozesses. Für jedes $\dot{\omega} \in \dot{\Omega}$ heißt die Abbildung $t \mapsto X_t(\dot{\omega})$ von T nach Ω ein Pfad des Prozesses.

Für unsere Gittergase bedeutet das:

$$X_t : \dot{\Omega} \rightarrow \Omega = E^S,$$

wobei ich die Menge der Gitterpunkte weiterhin mit S , eine abzählbar unendliche Menge, bezeichne und E sei die Menge $\{0, 1\}$. Die Parametermenge ist die Zeit T . Der Pfad ordnet an allen Punkten des Gitters der Zeit ebenfalls den Zustand 0 oder 1 zu:

$$X(\omega) : T \rightarrow \Omega.$$

Damit ist der Zustandsraum der obigen Definition $\Omega = E^S$ der bekannte Konfigurationsraum. Hinter $\dot{\Omega}$ versteckt sich Ω^T , was ich später erläutern werde.

Kommen wir nun zu einem speziellen Prozeß, dem Markov-Prozeß, siehe [6]. Wie bereits im ersten Kapitel erwähnt, werden uns die Wahrscheinlichkeitskerne dabei behilflich sein. Spezifizierungen sind Familien von W-Kernen, die räumlich indiziert werden. Nun benötigen wir durch die Zeit indizierte W-Kerne.

Definition 3.1.2. Es sei $(P_t)_{t \in T}$ eine durch T indizierte Familie von Wahrscheinlichkeitskernen auf einem Meßraum (Ω, \mathfrak{F}) . Gelten die Chapman-Kolmogoroff-Gleichungen

$$P_{s+t} = P_s P_t \quad \forall s, t \in T, \quad (3.1)$$

so heißt $(P_t)_{t \in T}$ eine Halbgruppe von Kernen auf Ω . Die Halbgruppe heißt Markovsche Halbgruppe.

3. Markov-Prozeß und zellulare Automaten

Beispiel 3.1.1. Denkt man an die Übergangsmatrizen P eines diskreten Markov-Prozesses

$$p(n+1) = Pp(n),$$

so erfüllen die Matrizen die Gleichung

$$P^{n+m} = P^n P^m \quad \forall n, m \in \mathbb{N}.$$

Bezeichnen wir mit $P(x, y)$ die Einträge der Matrix mit $x, y \in \mathbb{N}$ und mit A eine Teilmenge der Potenzmenge $\mathfrak{P}(\mathbb{N})$ von \mathbb{N} , dann ist durch

$$K(A|x) = \sum_{y \in A} P(x, y)$$

ein W-Kern definiert.

Nun will ich mit der Markovschen Halbgruppe Wahrscheinlichkeitsmaße definieren und benötige dafür folgende Definition.

Definition 3.1.3. Mit π_J^H sei die Projektion von Ω^H auf Ω^J mit $J \subset H$ und $J, H \in \mathfrak{h}(T)$ bezeichnet, wobei ich alle nichtleere, endliche Teilmengen von T mit $\mathfrak{h}(T)$ bezeichne. Des Weiteren gelte für die W-Maße P_H auf Ω^H und P_J auf Ω^J

$$P_J = \pi_J^H(P_H).$$

Dann heißt die Familie $(P_J)_{J \in \mathfrak{h}(T)}$ projektiv.

Es gilt der folgende Satz, den ich wie auch weitere folgende Sätze ohne Beweis liefern werde. Die Beweise findet man in [6].

Satz 3.1.1. Ist Ω ein polnischer Raum, d. h. ist Ω ein topologischer, bzgl. seiner Metrik vollständiger, Raum mit einer abzählbaren Basis, so existiert zu jeder projektiven Familie $(P_J)_{J \in \mathfrak{h}(T)}$ von W-Maßen auf $(\Omega^J, \mathfrak{F}^J)$ genau ein W-Maß P auf $(\Omega^T, \mathfrak{F}^T)$ mit

$$\pi_J(P) = P_J \quad \forall J \in \mathfrak{h}(T).$$

Bemerkung 3.1.1. $\Omega = \{0, 1\}^S$ ist ein polnischer Raum.

Kommen wir nun zurück zu den Markovschen Halbgruppen und dem Ziel, mit diesen W-Maße zu definieren.

Satz 3.1.2. Auf einem Meßraum (Ω, \mathfrak{F}) seien $(P_t)_{t \in T}$ eine Markovsche Halbgruppe und μ ein W-Maß. Für jede Menge $J = \{t_1, \dots, t_n\} \in \mathfrak{h}(T)$ mit Elementen $t_1 < t_2 < \dots < t_n$ und jedes $A \in \mathfrak{E}^J$ setze man

$$P_J := \int \int \dots \int 1_A(\omega_1, \dots, \omega_n) P_{t_n - t_{n-1}}(d\omega_n | \omega_{n-1}) \cdot \dots \cdot P_{t_1}(d\omega_1 | \omega_0) \mu(d\omega_0).$$

Dann ist $(P_J)_{J \in \mathfrak{h}(T)}$ eine projektive Familie von W-Maßen auf $(\Omega^J, \mathfrak{F}^J)$.

Nun reicht uns das W-Maß nicht aus, wir wollten mit den Markovschen Halbgruppen einen stochastischen Prozeß definieren. Der folgende Satz ermöglicht die Konstruktion eines Prozesses.

Satz 3.1.3. *Ist Ω ein polnischer Raum, \mathfrak{F} die σ -Algebra seiner Borelschen Mengen und T eine beliebige nichtleere Menge, so existiert zu jeder projektiven Familie $(P_J)_{J \in \mathfrak{h}(T)}$ von W-Maßen auf $(\Omega^J, \mathfrak{F}^J)$ ein stochastischer Prozeß mit Zustandsraum Ω und Parametermenge T derart, dass $(P_J)_{J \in \mathfrak{h}(T)}$ die Familie seiner endlich-dimensionalen Verteilungen ist.*

Definition 3.1.4. Sei X eine (Ω, \mathfrak{F}) -Zufallsvariable auf einem Wahrscheinlichkeitsraum $(\Omega, \mathfrak{F}, P)$. Dann heißt das Bildmaß

$$P_X := X(P)$$

die Verteilung von X bezüglich P . Es gilt $P_X(A) = P\{X \in A\} = P(X^{-1}(A))$ für $A \in \mathfrak{F}$.

Kommen wir auf **Satz 3.1.2** zurück. Ist Ω ein polnischer Raum und \mathfrak{F} die σ -Algebra seiner Borelschen Mengen, so erhalten wir mit **Satz 3.1.3**, dass die projektive Familie $(P_J)_{J \in \mathfrak{h}(T)}$ die Familie der endlich-dimensionalen Verteilungen des stochastischen Prozesses mit Zustandsraum Ω ist. Die Familie hängt bei gegebener Markovscher Halbgruppe nur noch vom Startmaß μ ab. Der zu (P_J) gehörige kanonische Prozeß ist

$$(\Omega^T, \mathfrak{F}^T, P^\mu, (X_t = \pi_{\{t\}})_{t \in T})$$

ist. Für die einpunktige Menge $J = \{t\} \subset T$ gilt

$$P_{\{t\}}(A) = \int_{\Omega} P_t(A|\omega_0) \mu(d\omega_0) = P^\mu\{X_t \in A\} \quad \forall A \in \mathfrak{F}.$$

3.2. Markov-Prozeß

Im Folgenden wähle ich für das Startmaß μ das durch die Einheitsmasse in ω definierte Maß ϵ_ω mit $\omega \in \Omega$ und schreibe für $P^\mu = P^{\epsilon_\omega} = P^\omega$. Folglich gilt mit (3.1)

$$P_t(A|\omega) = P^\omega\{X_t \in A\}. \quad (3.2)$$

$P^\omega\{X_t \in A\}$ ist also die Wahrscheinlichkeit für $X_t \in A$, wenn zur Zeit $t = 0$ das Ereignis ω eingetreten ist.

Definition 3.2.1. Ein stochastischer Prozeß mit total geordneter Parametermenge T besitzt die elementare Markov-Eigenschaft, wenn für jede Menge $A \in \mathfrak{F}$ und jedes Paar $s, t \in T$ mit $s < t$ gilt:

$$P^\omega\{X_t \in A | \mathfrak{F}_s\} = P^\omega\{X_t \in A | X_s\} \quad P^\omega\text{-fast sicher.}$$

Dabei ist \mathfrak{F}_s die von allen Zufallsvariablen X_r mit $r \leq s$ erzeugte σ -Algebra. Sie enthält die Informationen der s -Vergangenheit.

3. Markov-Prozeß und zellulare Automaten

Der nächste Satz ermöglicht uns eine günstige Schreibweise, so dass wir nach der Flut von Sätzen endlich den für uns wichtigen Markov-Prozeß definieren können.

Satz 3.2.1. Sei $(\Omega^{\mathbb{R}_+}, \mathfrak{F}^{\mathbb{R}_+}, P^\omega, (X_t)_{t \in \mathbb{R}_+})$ ein stochastischer Prozeß mit Zustandsraum (Ω, \mathfrak{F}) und Parametermenge \mathbb{R}_+ , dessen endlich-dimensionale Verteilungen sich gemäß **Satz 3.1.2** aus einer Markovschen Halbgruppe $(P_t)_{t \in \mathbb{R}_+}$ und einer Startwahrscheinlichkeit $\mu = \epsilon_\omega$ auf (Ω, \mathfrak{F}) ableiten. Dann besitzt der Prozeß die elementare Markov-Eigenschaft. Weiterhin gilt für beliebige $A \in \mathfrak{F}$ und $s, t \in \mathbb{R}_+$ mit $s < t$

$$P^x \{X_t \in A | \mathfrak{F}_s\} = P_{t-s}(A | X_s) \quad P^\omega\text{-fast sicher.} \quad (3.3)$$

Aus (3.3) folgt

$$P^\omega \{X_{s+t} \in A | \mathfrak{F}_s\} = P_t(A | X_s) \quad P^\omega\text{-fast sicher.}$$

Zusammen mit Gleichung (3.2) folgt

$$P^\omega \{X_{s+t} \in A | \mathfrak{F}_s\} = P^{X_s} \{X_t \in A\} \quad P^\omega\text{-fast sicher.}$$

Kommen wir nun zur für uns wesentlichen Definition des Markov-Prozesses.

Definition 3.2.2. Sei (Ω, \mathfrak{F}) ein Meßraum. Markovscher Prozeß mit Zustandsraum Ω heißt dann jedes Quadrupel $(\Omega^{\mathbb{R}_+}, \mathfrak{F}^{\mathbb{R}_+}, (P^\omega)_{\omega \in \Omega}, (X_t)_{t \in \mathbb{R}_+})$ mit folgenden Eigenschaften

- Für jedes $\omega \in \Omega$ ist $(\Omega^{\mathbb{R}_+}, \mathfrak{F}^{\mathbb{R}_+}, (P^\omega), (X_t)_{t \in \mathbb{R}_+})$ ein stochastischer Prozeß mit Zustandsraum Ω .
- Für jedes $B \in \mathfrak{F}^{\mathbb{R}_+}$ ist $\omega \mapsto P^\omega(B)$ \mathfrak{F} -messbar.
- $P^\omega \{X_{s+t} \in A | \mathfrak{F}_s\} = P^{X_s} \{X_t \in A\} \quad P^\omega\text{-fast sicher}$
für alle $s, t \in \mathbb{R}_+$, $\omega \in \Omega$ und $A \in \mathfrak{F}$.

Bemerkung 3.2.1. Anschaulich kann man sagen, dass beim Markov-Prozeß die Information über die s -Vergangenheit gleichwertig mit dem Zustand zur Zeit s ist.

Ist der Zustandsraum Ω endlich oder abzählbar, so spricht man von einer kontinuierlichen Markov-Kette. Im Falle einer zusätzlich noch abzählbaren Parametermenge, also z. B. \mathbb{N} statt \mathbb{R} , heißt der Markov-Prozeß Markov-Kette. Hängen die Übergangswahrscheinlichkeiten nicht vom Parameter ab, so spricht man von einer homogenen Markov-Kette.

Wir können nun einen zusammenfassenden für uns wichtigen Satz formulieren. Er stellt die Bedeutung der Markovschen Halbgruppen und damit die der Familien von W-Kernen heraus.

Satz 3.2.2. Zu jeder Markovschen Halbgruppe $(P_t)_{t \in \mathbb{R}_+}$ auf einem polnischen Raum Ω existiert ein Markov-Prozeß $(\Omega^{\mathbb{R}_+}, \mathfrak{F}^{\mathbb{R}_+}, (P^\omega)_{\omega \in \Omega}, (X_t)_{t \in \mathbb{R}_+})$ mit Zustandsraum Ω derart, dass für alle $t \in \mathbb{R}_+$, $\omega \in \Omega$ und $A \in \mathfrak{F}$ Gleichung (3.2) gilt:

$$P_t(A | \omega) = P^\omega \{X_t \in A\}.$$

Umgekehrt gilt, dass bei einem wie oben gegebenem Markov-Prozeß durch Gleichung (3.2) eine Markovsche Halbgruppe auf (Ω, \mathfrak{F}) definiert wird.

Ich habe jetzt die Definition und alle notwendigen Grundlagen für die folgenden Abschnitte gegeben, in denen ich durch die Zeit indizierte Markov-Prozesse einführen will, um bedingte Gibbs-Maße als stationäre Maße dieser Markov-Prozesse zu identifizieren.

3.3. Übergangswahrscheinlichkeiten kontinuierlicher Markov-Ketten

Sehen wir uns noch einmal die Chapman-Kolmogoroff-Gleichungen (3.1) an und schreiben diese aus:

$$P_{s+t}(A|\omega) = \int P_s(d\omega|\omega)P_t(A|\omega),$$

wobei ω ein Element des Zustandsraums Ω und A ein Element der σ -Algebra des Zustandsraums \mathfrak{F} ist. Mit s und t bezeichne ich Elemente der Parametermenge T , z. B. der Zeit. Ist der Zustandsraum abzählbar, so können wir

$$P_{s+t}(i, j) = \sum_k P_s(i, k)P_t(k, j) \quad (3.4)$$

schreiben. Dabei sind i, k, j Elemente des abzählbaren Zustandsraums E , s und t sind wieder Elemente der Parametermenge und $P_{s+t}(i, j)$ ist die Übergangswahrscheinlichkeit von Zustand i nach Zustand j in der Zeit $s + t$. Die Funktionen $P_t(i, j)$ seien stetig und differenzierbar für $t \geq 0$ mit

$$\left| \frac{dP_t(i, j)}{dt} \right| < K = \text{konst.} \quad \forall i, j \in E, t \geq 0.$$

Bemerkung 3.3.1. Wir berücksichtigen nun also nicht nur Übergangswahrscheinlichkeiten von Zustand i nach j in der Zeit von n nach $n + 1$, sondern lassen beliebige Zeitschritte zu. Im Fall des endlichen Zustandsraums haben wir also eine Markov-Kette mit kontinuierlicher Parametermenge.

Wir betrachten im Folgenden eine Markov-Kette mit kontinuierlicher Parametermenge und endlichem Zustandsraum E . Man definiert die Sprungrate $Q(i, j)$ von Zustand i zu Zustand j , vgl. [7, 8, 16], durch

$$q(i, j) = Q(i, j) = \left. \frac{dP_t(i, j)}{dt} \right|_{t=0} = \lim_{h \rightarrow 0} \frac{P_h(i, j) - P_0(i, j)}{h} \quad j \neq i. \quad (3.5)$$

Bemerkung 3.3.2. Wir werden später auch $(Q(i, i))_{i \in E}$ definieren und die Matrix $(Q(i, j))_{i, j \in E}$ als Markov-Generator bezeichnen. Die Funktion $P_t(i, j)$ hat bei 0 den Wert 1 oder 0. Ist $i \neq j$, so ist die Übergangswahrscheinlichkeit 0 für $t = 0$, da anschaulich für den Zustandswechsel keine Zeit bleibt. Dementsprechend gilt für $i = j$, dass $P_0(i, i) = 1$ ist, da der Zustand i für $t = 0$ nicht verlassen werden kann. Für die Markovsche Halbgruppe $P_t = (P_t(i, j))_{i, j \in E}$ gilt also $P_0 = I$. Da die Übergangswahrscheinlichkeiten zwischen 0 und 1 liegen gilt

$$Q(i, i) < 0 \quad \text{und} \quad Q(i, j) > 0, \quad \forall i, j \in E.$$

3. Markov-Prozeß und zellulare Automaten

An dieser Stelle möchte ich einen Satz zitieren, der unsere folgenden Rechnungen ermöglicht, für den Beweis vgl. [7].

Satz 3.3.1. *Es sei P_t eine stetige Markovsche Halbgruppe auf einem abzählbaren Zustandsraum E , es gilt also für alle $t, s \geq 0$*

- P_t ist eine stochastische Matrix.
- $P_0 = I$.
- $P_{t+s} = P_t P_s$.

Dann existiert für alle $i \in E$

$$\lambda_i := \lim_{h \searrow 0} \frac{1 - P_h(i, i)}{h} \in [0, \infty]$$

und für alle $i, j \in E$

$$q(i, j) = \lim_{h \searrow 0} \frac{P_h(i, j)}{h} \in [0, \infty).$$

Bemerkung 3.3.3. Da wir endliche Zustandsräume betrachten, gilt sogar $\lambda_i < \infty$.

Beweis. Es gilt $\sum_{j \in E} P_h(i, j) = 1$ und damit

$$\frac{1 - P_h(i, i)}{h} = \sum_{j \neq i} \frac{P_h(i, j)}{h}.$$

Nun bilden wir auf beiden Seiten den Grenzwert $\lim_{h \searrow 0}$. Auf der rechten Seite darf die Grenzwertbildung und die endliche Summation vertauscht werden und es gilt $q(i, j) < \infty$, damit folgt

$$\lambda_i = \lim_{h \searrow 0} \sum_{j \neq i} \frac{P_h(i, j)}{h} = \sum_{j \neq i} q(i, j) < \infty.$$

□

Die Halbgruppe P_t wird stabil genannt, wenn $\lambda_i < \infty$ gilt. Man bezeichnet sie als konservativ, falls $\lambda_i = \sum_{j \neq i} q(i, j)$. Im Falle des endlichen Zustandsraums haben wir also eine stabile und konservative Halbgruppe.

Beispiel 3.3.1. Sei $N(t)$ ein Poisson-Prozeß mit der Rate λ , wobei $N(t)$ z. B. die Anzahl von Sprüngen ist, die sich bis zur Zeit t ereignen. Die Rate λ gibt die durchschnittliche Anzahl von Sprüngen pro Zeiteinheit an. Die Wahrscheinlichkeit für $N(t) = n$ ist gegeben durch

$$P(N(t) = n) = \exp(-\lambda t) \frac{(\lambda t)^n}{n!}.$$

Des weiteren sei eine Markov-Kette mit der Familie von Zufallsvariablen $(Y_n)_{n \in \mathbb{N}}$ gegeben. Die Übergangswahrscheinlichkeit vom Zustand i in den Zustand j in einem Zeitschritt der Markov-Kette sei durch $u(i, j)$ gegeben.

3.3. Übergangswahrscheinlichkeiten kontinuierlicher Markov-Ketten

Wir erhalten nun eine Markov-Kette mit kontinuierlicher Parametermenge, indem wir folgende Familie von Zufallsvariablen betrachten: $X_t = Y_{N(t)}$. Wir erhalten nun für jede Zeit $t > 0$ eine Übergangswahrscheinlichkeit

$$P(X_t = j | X_0 = i) = P_t(i, j) = \sum_{n=0}^{\infty} \exp(-\lambda t) \frac{(\lambda t)^n}{n!} u^n(i, j). \quad (3.6)$$

Kümmern wir uns nun um die Sprungrate. Die Wahrscheinlichkeit, dass mindestens zwei Sprünge in der Zeit h stattfinden, ergibt sich zu

$$\begin{aligned} 1 - (\exp(-\lambda h) + \lambda h \exp(-\lambda h)) &= 1 - (1 + \lambda h) \left(1 - \lambda h + \frac{(\lambda h)^2}{2!} + \dots \right) \\ &= \frac{(\lambda h)^2}{2!} + \dots \end{aligned}$$

Dividieren wir diese Wahrscheinlichkeit durch h und lassen h gegen 0 streben, so konvergiert der Quotient ebenfalls gegen Null. Die Übergangswahrscheinlichkeit von i nach j in Null Sprüngen ist Null, $u^0(i, j) = 0$ für $i \neq j$. Für die Übergangswahrscheinlichkeit (3.6) bleibt der Summand für $n = 1$

$$\frac{P_h(i, j)}{h} \approx \lambda \exp(-\lambda h) u(i, j) \rightarrow \lambda u(i, j) \text{ für } h \rightarrow 0.$$

Die Sprungrate lautet also $q(i, j) = \lambda u(i, j)$ und bedeutet, dass der Zustand i mit der Rate λ verlassen wird und der Zustand j mit der Wahrscheinlichkeit $u(i, j)$ angenommen wird.

Mit Hilfe des Beispiels kann allgemein gesagt werden, dass $q(i, j)$ die Rate der Sprünge von i nach j ist.

Nun wollen wir uns dem Fall zuwenden, dass wir die Sprungrate $q(i, j)$, nicht aber die Übergangswahrscheinlichkeit $P_t(i, j)$ kennen. Unser Ziel wird es sein, mit der Sprungrate die Übergangswahrscheinlichkeit zu berechnen. Dazu nutze ich die Chapman-Kolmogorov-Gleichungen (3.4). Doch zuvor möchte ich die Rate λ_i , mit der der Zustand i verlassen wird, definieren als

$$\lambda_i = \sum_{j \neq i} q(i, j).$$

Die Wahrscheinlichkeit $r(i, j)$, dass der Zustand i verlassen und der Zustand j angenommen wird, ergibt sich zu

$$r(i, j) = \frac{q(i, j)}{\lambda_i}.$$

Zurück zu unserem Ziel, die Übergangswahrscheinlichkeit zu bestimmen. Diese werden wir mit Hilfe der Ableitung der Übergangswahrscheinlichkeit charakterisieren

$$P'_t(i, j) = \lim_{h \rightarrow 0} \frac{P_{t+h}(i, j) - P_t(i, j)}{h}.$$

Dabei hilft uns, wie bereits erwähnt, Gleichung (3.4)

$$\begin{aligned} P_{t+h}(i, j) - P_t(i, j) &= \sum_k P_h(i, k) P_t(k, j) - P_t(i, j) \\ &= \sum_{k \neq i} P_h(i, k) P_t(k, j) + (P_h(i, i) - 1) P_t(i, j). \end{aligned} \quad (3.7)$$

3. Markov-Prozeß und zellulare Automaten

Offensichtlich müssen wir nun die rechte Seite der Gleichung (3.7) durch h dividieren und den Grenzwert für $h \rightarrow 0$ bilden, um die Ableitung der Übergangswahrscheinlichkeit zu berechnen. Betrachten wir zunächst die Summe der rechten Seite der Gleichung (3.7) und nutzen die Definition der Sprungrate (3.5). Durch Vertauschen der Summe und des Grenzprozesses erhalten wir

$$\lim_{h \rightarrow 0} \frac{1}{h} \sum_{k \neq i} P_h(i, k) P_t(k, j) = \sum_{k \neq i} q(i, k) P_t(k, j). \quad (3.8)$$

Mit $1 - P_h(i, i) = \sum_{k \neq i} P_h(i, k)$ folgt

$$\lim_{h \rightarrow 0} \frac{P_h(i, i) - 1}{h} = - \lim_{h \rightarrow 0} \sum_{k \neq i} \frac{P_h(i, k)}{h} = - \sum_{k \neq i} q(i, k) = -\lambda_i.$$

Damit erhalten wir

$$\lim_{h \rightarrow 0} \frac{P_h(i, i) - 1}{h} P_t(i, j) = -\lambda_i P_t(i, j). \quad (3.9)$$

Dividieren wir nun beide Seiten der Gleichung (3.7) durch h und bilden den Grenzwert $h \rightarrow 0$ ergibt sich unter Ausnutzung der Gleichungen (3.8) und (3.9)

$$P'_t(i, j) = \sum_{k \neq i} q(i, k) P_t(k, j) - \lambda_i P_t(i, j). \quad (3.10)$$

Die Definition der Matrix Q

$$Q(i, j) = \begin{cases} q(i, j) & \text{für } j \neq i \\ -\lambda_i & \text{für } j = i \end{cases} \quad (3.11)$$

erlaubt eine Vereinfachung der Gleichung (3.10) zu

$$P'_t = Q P_t. \quad (3.12)$$

Bemerkung 3.3.4. P'_t und P_t sind Matrizen! Die Differentialgleichung (3.12) ist mit der Anfangsbedingung $P_0 = I$ eindeutig lösbar.

Wir haben nun mit (3.12) eine Bestimmungsgleichung für die Übergangswahrscheinlichkeiten. Setzen wir $P_t = \exp(Qt)$ und definieren wir

$$\exp(Qt) = \sum_{n=0}^{\infty} \frac{(Qt)^n}{n!} = \sum_{n=0}^{\infty} Q^n \frac{t^n}{n!},$$

dann zeigt

$$\frac{d}{dt} \exp(Qt) = \sum_{n=0}^{\infty} Q^n \frac{t^{n-1}}{(n-1)!} = \sum_{n=0}^{\infty} Q Q^{n-1} \frac{t^{n-1}}{(n-1)!} = Q \exp(Qt),$$

3.4. Stationäre Maße kontinuierlicher Markov-Ketten

dass $P_t = \exp(Qt)$ Gleichung (3.12) erfüllt. Gleichung (3.12) wird als Kolmogorovs *backward equation* bezeichnet. Kolmogorovs *forward equation*

$$P'_t = P_t Q \quad (3.13)$$

erhält man mit folgendem Ansatz

$$\begin{aligned} P_{t+h}(i, j) - P_t(i, j) &= \sum_k P_t(i, k) P_h(k, j) - P_t(i, j) \\ &= \sum_{k \neq i} P_t(i, k) P_h(k, j) + (P_h(i, i) - 1) P_t(i, j). \end{aligned}$$

Die Gleichungen (3.12) und (3.13) zeigen, dass

$$P_t Q = Q P_t$$

gilt.

Die Matrix $\exp(Qt)$ besteht aus den Übergangswahrscheinlichkeiten bzw. den bedingten Wahrscheinlichkeiten und bildet die Markovsche Halbgruppe, wobei Q der so genannte Markov-Generator ist. Der Infinitesimalgenerator generiert die Markovsche Halbgruppe und damit das Wahrscheinlichkeitsmaß des Prozesses. Sehen wir uns den Generator genauer an. Es gilt nach den vorherigen Überlegungen

$$P'_t = Q P_t.$$

Des Weiteren ist eine Ableitung über den Grenzwert des Differenzenquotienten definiert und mit der Halbgruppeneigenschaft läßt sich folgende Umformung vornehmen

$$P'_t = \lim_{h \rightarrow 0} \frac{P_{t+h} - P_t}{h} = \lim_{h \rightarrow 0} P_t \frac{P_h - I}{h}.$$

Für den Markov-Generator Q ergibt sich durch einen Vergleich der letzten beiden Gleichungen

$$Q = \lim_{h \rightarrow 0} \frac{P_h - I}{h}.$$

3.4. Stationäre Maße kontinuierlicher Markov-Ketten

Im Fall der kontinuierlichen Parametermenge ist die Übergangsmatrix P_t zeitabhängig. Für eine stationäre Verteilung π muss

$$\pi P_t = \pi \quad \forall t > 0 \quad (3.14)$$

gelten.

Bemerkung 3.4.1. Durch π ist ein Vektor gegeben und $\pi(i)$ ist die stationäre Wahrscheinlichkeit für den Zustand i . Die stationäre Verteilung muss nicht eindeutig sein, siehe **Bemerkung 3.4.4. Satz 4.1.1** formuliert Bedingungen für die Existenz einer eindeutigen stationären Verteilung einer Markov-Kette mit diskreter Parametermenge.

Die Summe der Zeileneinträge der Matrix P_t ergibt 1.

3. Markov-Prozeß und zellulare Automaten

Auf Grund der im letzten Abschnitt herausgearbeiteten Schwierigkeiten mit der Berechnung von P_t lohnt es sich, ein weiteres Kriterium für stationäre Verteilungen kontinuierlicher Markov-Ketten einzuführen.

Satz 3.4.1. *Durch π ist genau dann eine stationäre Verteilung gegeben, wenn*

$$\pi Q = 0. \quad (3.15)$$

Bevor wir den Satz beweisen, folgt eine

Bemerkung 3.4.2. Setzen wir in die Gleichung (3.15) die Definition der Matrix Q (3.11) ein, so erhalten wir

$$\sum_{k \neq j} \pi(k) q(k, j) = \pi(j) \lambda_j. \quad (3.16)$$

Es sei λ_j die Rate, mit der der Zustand j verlassen wird, während $q(k, j)$ die Rate angibt, mit der vom Zustand k zum Zustand j gewechselt wird. Also steht links in der Gleichung (3.16) die Rate, mit der der Zustand j angenommen, und rechts die Rate, mit der der Zustand j verlassen wird. Anschaulich formuliert bedeutet dies: „Zufluß = Abfluß“. Dies ist tatsächlich die anschauliche Interpretation einer stationären Verteilung.

Kommen wir nun zum Beweis von **Satz 3.4.1**.

Beweis. Beginnen wir zu zeigen, dass, wenn eine Verteilung π stationär ist, $\pi Q = 0$. Multiplizieren wir Gleichung (3.12) mit $\pi(i)$ und summieren anschließend, dann erhalten wir

$$\sum_i \pi(i) P'_t(i, j) = \sum_{i, k} \pi(i) P_t(i, k) Q(k, j). \quad (3.17)$$

Ziehen wir nun die Ableitung aus der Summe und nutzen (3.14), bekommen wir für die linke Seite

$$\frac{d}{dt} \left(\sum_i \pi(i) P_t(i, j) \right) = \frac{d}{dt} \pi(j) = 0.$$

Nun führen wir auf der rechten Seite der Gleichung (3.17) die Summation über i aus:

$$\sum_i \pi(i) P_t(i, k) = \pi(k).$$

Jetzt resultiert aus Gleichung (3.17) die Beziehung

$$0 = \sum_k \pi(k) Q(k, j) = (\pi Q)_j.$$

3.5. Bedingte Gibbs-Maße als stationäre Maße einer kontinuierlichen Markov-Kette

Dabei ist $(\pi Q)_j$ die j -te Komponente des Vektors πQ .

Nun folgern wir aus $\pi Q = 0$, dass π eine stationäre Verteilung ist. Dazu multiplizieren wir Gleichung (3.12) mit π und summieren:

$$\sum_i \pi(i) P'_t(i, j) = \sum_k (\pi Q)_k P_t(k, j).$$

Nutzen wir die Voraussetzung $\pi Q = 0$, steht auf der rechten Seite Null. Auf der linken Seite vertauschen wir die Ableitung und die Summe, so dass sich

$$\frac{d}{dt} \sum_i \pi(i) P_t(i, j) = 0$$

ergibt. Es ist πP_t also konstant und muss überall den Wert πP_0 annehmen. Dieser Wert ist π , da in dem Zeitintervall von Null Zeiteinheiten kein Übergang statt findet. \square

Die folgende Definition werden wir später benötigen. Sie ist stärker als die Definition einer stationären Verteilung.

Definition 3.4.1. Eine Verteilung π heißt reversibel, wenn

$$\pi(k)q(k, j) = \pi(j)q(j, k) \quad \forall j \neq k.$$

Bemerkung 3.4.3. Jede reversible Verteilung ist stationär, da aus

$$\pi(k)q(k, j) = \pi(j)q(j, k) \quad \forall j \neq k$$

auch

$$\sum_{k \neq j} \pi(k)q(k, j) - \pi(j)q(j, k) = 0$$

folgt.

Bemerkung 3.4.4. Ist eine Markov-Kette irreduzibel, d. h. ist $P_t(i, j) > 0$, $\forall i, j \in E$ für jeweils ein $t \in T$, und hat das System (3.15) eine eindeutige Lösung π mit $\sum \pi(i) = 1$ und $\pi(i) > 0$, $\forall i \in E$, dann gilt $\lim_{t \rightarrow \infty} P_t(i, j) = \pi(j)$. Ist der Zustandsraum endlich, so existiert für irreduzible Ketten immer eine eindeutige Lösung π mit $\sum \pi(i) = 1$ und $\pi(i) > 0$, $\forall i \in E$. Ist der Zustandsraum abzählbar, so kann es passieren, dass $\lim_{t \rightarrow \infty} P_t(i, j) = 0$ gilt und das System (3.15) nur die triviale Lösung $\pi = (0, 0, 0, \dots)$ besitzt, siehe [1, 8, 33].

3.5. Bedingte Gibbs-Maße als stationäre Maße einer kontinuierlichen Markov-Kette

Wir betrachten nun wieder unser Gitter $S = \mathbb{Z}^d$ und den Zustandsraum $E = \{0, 1\}$. Den Zustand (besetzt oder unbesetzt) an der Stelle i geben wir mit σ_i , $\sigma \in \Omega = E^S$ an. Als nächstes definiere ich folgende Sprungrate für alle $i \in S$:

$$c(i, \sigma) = \exp \left(\sum_{A \ni i} \Phi(A) \sigma(A) \right) \quad \text{mit } A \in \mathcal{S},$$

3. Markov-Prozeß und zellulare Automaten

wobei Φ ein translationsinvariantes Potential mit beschränkter Reichweite ist und $\sigma(A) = \prod_{j \in A} \sigma_j$, vgl. [15].

Bemerkung 3.5.1. Die Optik der Sprungrate c ist eine andere als die der Sprungrate q in den vorherigen Abschnitten. Tatsächlich meinen aber beide das Gleiche. $q(i, j)$ war die Rate von Zustand i zu Zustand j . Durch $c(i, \sigma)$ ist die Rate definiert, dass sich der Zustand der gegebenen Konfiguration $\sigma \in \Omega$ an der Stelle $i \in S$ ändert und damit die Konfiguration $\sigma^i \in \Omega$ angenommen wird. Der neue Zustand wird bei $c(i, \sigma)$ nicht angegeben, da dieser auf Grund der nur zwei möglichen Zustände $E = \{0, 1\}$ eindeutig durch die Angabe des alten Zustands definiert ist.

Da sich nur der Zustand eines Gitterpunktes $i \in S$ verändert können die alten Zustände $\sigma \in \Omega$ nicht in beliebige neue Zustände wechseln. Durch eine Verkettung von Sprüngen ist allerdings jeder Zustand erreichbar.

Wie wir es bereits kennen, teile ich S in $\Lambda \in \mathcal{S}$ und $S \setminus \Lambda$ auf. Es sei $\zeta \in E^\Lambda$ und $\eta \in E^{S \setminus \Lambda}$. Indem ich die Sprungrate auf Λ reduziere

$$c_\Lambda(i, \sigma) = c(i, \sigma) 1_\Lambda(i) \text{ mit } \sigma \in E^S,$$

erhalte ich eine kontinuierliche Markov-Kette mit dem Zustandsraum E^Λ . Haben wir den Zustandsraum E^S , so handelt es sich nicht um eine Markov-Kette sondern einen Markov-Prozeß, da überabzählbar viele Ereignisse auftreten können. Der entscheidende Punkt des gesamten Kapitels ist, dass eine uns bekannte Familie von bedingten Wahrscheinlichkeiten, die Gibbssche Spezifizierung,

$$\gamma_\Lambda^\Phi(\sigma|\omega) = Z_\Lambda^\Phi(\omega)^{-1} \exp \left(- \sum_{A \cap \Lambda \neq \emptyset} \Phi(A) \zeta(A) \eta(A) \right) \text{ mit } \sigma_\Lambda = \zeta \text{ und } \omega_{S \setminus \Lambda} = \eta,$$

wobei $\zeta(A) = \prod_{i \in A, i \in \Lambda} \zeta_i$ und $\eta(A) = \prod_{j \in A, j \in S \setminus \Lambda} \eta_j$, ein stationäres Maß der kontinuierlichen Markov-Kette ist. Dies werde ich nun ausführen.

Wie oben bezeichnet σ^i die Konfiguration $\sigma \in \Omega$, wobei der Zustand an der Stelle i verändert wurde. Sehen wir uns einmal genauer an, welche Bedeutung $c_\Lambda(i, \sigma)$ besitzt. Es sind nur noch $i \in \Lambda$ erlaubt. Da nur über die Mengen A summiert werden darf, die i enthalten, gilt $A \cap \Lambda \neq \emptyset$. Auch für die Mengen A , über die in der Gibbsschen Spezifizierung summiert wird, gilt $A \cap \Lambda \neq \emptyset$. Im Unterschied zur Sprungrate muss A jedoch nicht i enthalten. Bilden wir den Quotienten

$$\frac{\gamma_\Lambda^\Phi(\sigma|\omega)}{\gamma_\Lambda^\Phi(\sigma^i|\omega)},$$

so läßt sich dieser mit $\exp \left(- \sum_{A \cap \Lambda \neq \emptyset, i \notin A} \Phi(A) \zeta(A) \eta(A) \right)$ kürzen, da sich hier für σ^i im Vergleich zu σ nichts ändert. Der Quotient läßt sich also schreiben als

$$\frac{\gamma_\Lambda^\Phi(\sigma|\omega)}{\gamma_\Lambda^\Phi(\sigma^i|\omega)} = \frac{\exp \left(- \sum_{A \ni i} \Phi(A) \zeta(A) \eta(A) \right)}{\exp \left(- \sum_{A \ni i} \Phi(A) \zeta^i(A) \eta(A) \right)} = \frac{\exp \left(\sum_{A \ni i} \Phi(A) \zeta^i(A) \eta(A) \right)}{\exp \left(\sum_{A \ni i} \Phi(A) \zeta(A) \eta(A) \right)}.$$

3.5. Bedingte Gibbs-Maße als stationäre Maße einer kontinuierlichen Markov-Kette

Dies ist aber gleich dem Quotienten

$$\frac{c_\Lambda(i, \sigma^i)}{c_\Lambda(i, \sigma)}.$$

Es gilt also

$$\frac{c_\Lambda(i, \sigma^i)}{c_\Lambda(i, \sigma)} = \frac{\exp\left(\sum_{A \ni i} \Phi(A) \zeta^i(A) \eta(A)\right)}{\exp\left(\sum_{A \ni i} \Phi(A) \zeta(A) \eta(A)\right)} = \frac{\gamma_\Lambda^\Phi(\sigma|\omega)}{\gamma_\Lambda^\Phi(\sigma^i|\omega)}$$

bzw.

$$c_\Lambda(i, \sigma^i) \gamma_\Lambda^\Phi(\sigma^i|\omega) = c_\Lambda(i, \sigma) \gamma_\Lambda^\Phi(\sigma|\omega). \quad (3.18)$$

Jetzt definiere ich die aus den vorherigen Abschnitten bekannte Matrix Q für die Markov-Kette mit Zustandsraum E^Λ

$$Q(\sigma, \tilde{\sigma}) = \begin{cases} c_\Lambda(i, \sigma) & \text{wenn } \tilde{\sigma} = \sigma^i, i \in \Lambda \\ -\sum_{i \in \Lambda} c_\Lambda(i, \sigma) & \text{wenn } \tilde{\sigma} = \sigma \\ 0 & \text{sonst} \end{cases}.$$

Der erste Fall ist äquivalent zu $\tilde{\sigma} \neq \sigma$. Nach **Satz 3.4.1** gilt für ein stationäres Maß μ

$$\sum_{\tilde{\sigma}} \mu(\tilde{\sigma}) Q(\tilde{\sigma}, \sigma) = 0 \quad \forall \sigma \in E^\Lambda.$$

Mit der Definition der Matrix Q und mit Gleichung (3.18) erhalten wir

$$\sum_{\tilde{\sigma}} \gamma_\Lambda^\Phi(\tilde{\sigma}|\omega) Q(\tilde{\sigma}, \sigma) = \gamma_\Lambda^\Phi(\sigma|\omega) \left(-\sum_{i \in \Lambda} c_\Lambda(i, \sigma) \right) + \sum_{i \in \Lambda} c_\Lambda(i, \sigma^i) \gamma_\Lambda^\Phi(\sigma^i|\omega) = 0.$$

Damit ist die Gibbssche Spezifizierung wirklich ein stationäres Maß der Markov-Kette mit Sprungrate $c(i, \sigma)$ und Zustandsraum E^Λ .

Bemerkung 3.5.2. Mit $\mathcal{G}_\Lambda(\Phi)$ sei die konvexe, abgeschlossene Hülle von $\{\gamma_\Lambda^\Phi|\omega_{S \setminus \Lambda} \in \cdot\}$. Damit haben wir Maße auf Ω , genauer sei \mathcal{G}_Λ die Gleichgewichtsverteilung in Λ bei vorgegebenen Randbedingungen $\omega_{S \setminus \Lambda}$. Nun kann gezeigt werden, dass für $\Lambda_k \rightarrow \mathbb{Z}^d$, $\gamma_{\Lambda_k}^\Phi \in \mathcal{G}_{\Lambda_k}(\Phi)$ für alle k und $\gamma_{\Lambda_k}^\Phi \rightarrow \mu$, das Maß μ eine stationäre Verteilung für den Prozeß auf Ω ist. Es gilt $\mu \in \mathcal{G}(\Phi)$, vgl. [15], [50].

Bemerkung 3.5.3. Dieser Markov-Prozeß erhält nicht die Anzahl der Teilchen in Λ , da jeder Zustand zulässig ist. Verändert man die Sprungrate zum Beispiel so, dass jeweils zwei Gitterpunkte $i, j \in \Lambda$ den Zustand tauschen, kann die Teilchenzahl erhalten werden. Die Sprungrate könnte für $\sigma_i \neq \sigma_j$ wie folgt aussehen

$$c(i, j, \sigma) = \exp \left(\sum_{A \ni i, A \not\ni j} \Phi(A) \sigma(A) + \sum_{A \ni j, A \not\ni i} \Phi(A) \sigma(A) + \sum_{A \ni i, j} \Phi(A) \sigma(A) \right) \text{ mit } A \in \mathcal{S}$$

3. Markov-Prozeß und zellulare Automaten

und muss für $\sigma_i = \sigma_j$ Null sein. Die Konfiguration $\sigma_h^{i,j}$ an der Stelle h nach dem Tausch der Zustände an den Stelle i, j ist gegeben durch

$$\sigma_h^{i,j} = \begin{cases} \sigma_j, & \text{wenn } h = i \\ \sigma_i, & \text{wenn } h = j \\ \sigma_h, & \text{wenn } h \neq i, j. \end{cases}$$

Damit die Gibbssche Spezifizierung nicht nur ein stationäres Maß sondern ein stationäres Wahrscheinlichkeitsmaß ist, muss die Normierungskonstante angepasst werden, sie darf nur noch die Zustände mit der festgelegten Teilchenzahl enthalten. Die durch diese Spezifizierung gekennzeichneten Gibbs-Maße bezeichnet man als *kanonische Gibbs-Maße*, siehe [25].

3.6. PCA und ESM

Nachdem wir nun mit Hilfe von Generatoren bedingte Gibbs-Maße als stationäre Maße kontinuierlicher Markov-Ketten kennengelernt haben, will ich in diesem Kapitel stochastische zellulare Automaten (Probabilistic Cellular Automata, PCA) betrachten. PCAs stellen Markov-Prozesse dar, deren Zustandsraum die Konfigurationen auf dem \mathbb{Z}^d sind. Die Zustände aller $i \in \mathbb{Z}^d$ verändern sich gleichzeitig und unabhängig, was für Simulationen ein entscheidender Unterschied zu der im vorherigen Abschnitt dargestellten Entwicklung und zu den Gibbs-Samplern ist, vgl. [7]. Als Parametermenge wird die diskrete Zeit gewählt. Die Generatoren stehen uns also nicht zur Verfügung. Dennoch möchte ich eine Verbindung zu den Gibbs-Maßen herstellen. Dies wird uns mit der Betrachtung von statistischen Gleichgewichtsmodellen (ESM) in $d + 1$ Dimensionen, d Dimensionen des Gitters plus die Zeit als weitere Dimension, gelingen.

Bemerkung 3.6.1. Betrachten wir Konfigurationen auf einer endlichen Teilmenge Λ von \mathbb{Z}^d , so ist der Zustandsraum $\Omega = E^\Lambda$ für ein endliches E endlich und der PCA ist eine Markov-Kette.

Wir werden allerdings $E^{\mathbb{Z}^d}$ betrachten und somit haben wir keinen abzählbaren Zustandsraum. Damit ist der PCA ein Markov-Prozeß.

Beginnen wir mit den stochastischen zellularen Automaten, siehe [49]. Wie bereits oben erwähnt handelt es sich um Markov-Ketten mit dem Zustandsraum $E^{\mathbb{Z}^d} = \{-1, +1\}^{\mathbb{Z}^d}$, wobei die Spins -1 und $+1$ in Anlehnung an das Ising-Modell gewählt wurden. Ich bezeichne mit $\sigma_{n,i} = \pm 1$ den Spin an der Stelle $i \in \mathbb{Z}^d$ zur Zeit $n \in \mathbb{Z}$ und mit $\sigma_n = \{\sigma_{n,i}\}_{i \in \mathbb{Z}^d}$ eine Konfiguration zur Zeit n . Zellulare Automaten verändern im Gegensatz zu den „Generator-Modellen“ zeitgleich und unabhängig an allen Stellen $i \in \mathbb{Z}^d$ den Spin. Die Übergangswahrscheinlichkeit für einen Spin an der Stelle i zur Zeit n hängt von der Konfiguration σ_{n-1} eingeschränkt auf eine endliche Nachbarschaft $N(i)$ von $i \in \mathbb{Z}^d$ zur Zeit $n - 1$ ab:

$$p_i(\sigma_{n,i} | \sigma_{n-1}) = p_i(\sigma_{n,i} | \sigma_{n-1, N(i)}).$$

Das bedingte Wahrscheinlichkeitsmaß für σ_n ist ein Produktmaß, welches mit $P(d\sigma_n|\sigma_{n-1})$ bezeichnet wird. Die bedingte Wahrscheinlichkeit für die Konfiguration σ_n erhält man formal als Produkt

$$P(d\sigma_{n+1}|\sigma_n) = \prod_{i \in \mathbb{Z}^d} p_i(d\sigma_{n,i}|\sigma_{n-1}). \quad (3.19)$$

Für die Übergangswahrscheinlichkeiten gilt

$$\sum_{\sigma_{n,i}=\pm 1} (\sigma_{n,i}|\sigma_{n-1}) = 1.$$

In den für uns interessanten Fällen hängt die Übergangswahrscheinlichkeit p_i nicht von der kompletten Konfiguration σ_{n-1} , sondern von den Nachbarspins von $\sigma_{n-1,i}$ ab. Im nächsten Abschnitt werde ich näher auf p_i eingehen und ein Beispiel liefern. Wie in früheren Abschnitten beschrieben liefert uns (3.19) für ein gegebenes Wahrscheinlichkeitsmaß μ_{n-1} auf der Konfiguration σ_{n-1} ein Wahrscheinlichkeitsmaß $\mu_n = \mu_{n-1}P$ auf der Konfiguration σ_n :

$$\mu_n(d\sigma_n) = \int \mu_{n-1}(d\sigma_{n-1})P(d\sigma_n|\sigma_{n-1}).$$

Ein Maß ν ist stationär, wenn $\nu = \nu P$, und periodisch mit Periode k , wenn $\nu P^k = \nu$ gilt.

Bemerkung 3.6.2. Wir haben bereits stationäre Wahrscheinlichkeitsverteilungen von kontinuierlichen Markov-Ketten betrachtet, vgl. (3.14), und werden dies noch für Markov-Ketten mit diskreter Parametermenge tun, siehe 4.1.1.

Als statistische Gleichgewichtsmodelle (Equilibrium statistical models, ESM) bezeichne ich die bereits ausführlich erklärten Gleichgewichtszustände, beschrieben durch die zeitunabhängigen, bedingten Gibbs-Maße. Jetzt betrachten wir allerdings ein $(d+1)$ -dimensionales Gitter, neben den n -Raumdimensionen kommt die Zeit als weitere Dimension hinzu. Eine Konfiguration σ auf $\{-1, +1\}^{\mathbb{Z}^{d+1}}$ wird beschrieben durch $\{\sigma_n\}_{n \in \mathbb{Z}}$. Betrachten wir nun in der PCA-Welt ein Maß ν , welches auf $\cup_{n \geq N} \mathbb{Z}_n^d$ mit $n \geq N, n \in \mathbb{Z}$ ab dem Zeitpunkt N stationär ist. Für $N \rightarrow \infty$ erhält man ein Maß μ_ν in der ESM-Welt, also auf $\{-1, +1\}^{\mathbb{Z}^{d+1}}$, welches in Zeit-Richtung translationsinvariant ist. Die Projektion des Maßes μ_ν auf ein beliebiges \mathbb{Z}_n^d ergibt das in der PCA-Welt stationäre Maß ν . Das entscheidende ist nun, sind die Übergangswahrscheinlichkeiten $p_i(\sigma_{n,i}|\sigma_{n-1})$ des zellularen Automaten echt positiv, so ist μ_ν ein Gibbs-Maß des ESM, vgl. [49]. Wichtig für die ESM-Welt ist die Zeitunabhängigkeit und die Hamiltonfunktion. Die Zeit haben wir in unser Modell in den $d+1$ Dimensionen integriert. Für die Hamiltonfunktion benötigen wir ein Potential, das aufgrund der Interaktion der Teilchen vom Spin einer Stelle x und von den Nachbarspins abhängt. Als Nachbarschaft einer Stelle $x = (n, i) \in \mathbb{Z}^{d+1}$ werden wir \mathbb{Z}_{n-1}^d ansehen. Unser Potential hängt also, wenn auch durch die $d+1$ Dimensionen verschleiert, wie die bedingten Wahrscheinlichkeiten des PCA vom vorherigen Zeitschritt ab. Wir definieren für $x = (n, i) \in \mathbb{Z}^{d+1}$

$$\exp(-\Phi_x(\sigma_x, \sigma_{n-1})) = p_i(\sigma_{n,i}|\sigma_{n-1}),$$

3. Markov-Prozeß und zellulare Automaten

siehe (2.3.2). Damit verändert sich die Normierungsbedingung zu

$$\sum_{\sigma_x = \pm 1} \exp(-\Phi_x(\sigma_x, \sigma_{n-1})) = 1.$$

Für die Hamiltonfunktion ergibt sich formal die Reihe

$$H_x^\Phi(\sigma) = \sum_{x=(n,i) \in \mathbb{Z}^{d+1}} \Phi_x(\sigma_x, \sigma_{n-1}).$$

3.7. Beispiel: Vom PCA zum ESM

Wir betrachten einen zellularen Automaten auf $\{-1, +1\}^{\mathbb{Z}}$, vgl. [24]. Der Spin $\sigma_{n,i}$ an der Stelle $i \in \mathbb{Z}$ zur Zeit $t \in \mathbb{Z}$ kann also den Wert $+1$ oder -1 annehmen. Bei geraden (ungeraden) Zeitschritten sollen sich die gerade (ungerade) indizierten Spins verändern. Die Veränderung des Spins σ_i soll in unserem Beispiel nur von den direkten Nachbarn abhängen, $p_i(\sigma_{n,i}|\sigma_{n-1}) = p_i(\sigma_{n,i}|\sigma_{n-1,i-1}, \sigma_{n-1,i+1})$. Mit der Regel

$$p_i(-\sigma_{n,i}|\sigma_{n-1,i-1}, \sigma_{n-1,i+1}) = 1 - p_i(\sigma_{n,i}|\sigma_{n-1,i-1}, \sigma_{n-1,i+1}) \quad (3.20)$$

ergeben sich vier den Automaten beschreibende Übergangswahrscheinlichkeiten

$$x = p(1|-1, -1), \quad z = p(1|1, 1), \quad y_1 = p(1|-1, 1), \quad y_2 = p(1|1, -1), \quad 0 \leq x, y_1, y_2, z \leq 1. \quad (3.21)$$

Mit Hilfe dieser Übergangswahrscheinlichkeiten wollen wir nun eine allgemeine Formel entwickeln

$$\begin{aligned} p_i(\sigma_{n,i}|\sigma_{n-1,i-1}, \sigma_{n-1,i+1}) &= p_i(\sigma_i|\sigma_{i-1}, \sigma_{i+1}) \\ &= \sigma_i T' + \sigma_i \sigma_{i+1} U'_1 + \sigma_i \sigma_{i-1} U'_2 + \sigma_i \sigma_{i+1} \sigma_{i-1} V' + c, \end{aligned} \quad (3.22)$$

wobei T, U_1, U_2 und V von (3.21) abhängen. Es muss gelten

$$x = T' - U'_1 - U'_2 + V', \quad z = T' + U'_1 + U'_2 + V', \quad y_1 = T' + U'_1 - U'_2 - V', \quad y_2 = T' - U'_1 + U'_2 - V'.$$

Wir erhalten

$$\begin{aligned} T' &= \frac{1}{4}x + \frac{1}{4}y_1 + \frac{1}{4}y_2 + \frac{1}{4}z, \\ U'_1 &= -\frac{1}{4}x + \frac{1}{4}y_1 - \frac{1}{4}y_2 + \frac{1}{4}z, \\ U'_2 &= -\frac{1}{4}x - \frac{1}{4}y_1 + \frac{1}{4}y_2 + \frac{1}{4}z, \\ V' &= \frac{1}{4}x - \frac{1}{4}y_1 - \frac{1}{4}y_2 + \frac{1}{4}z. \end{aligned}$$

Nun soll noch (3.20) erfüllt werden. Dafür subtrahieren wir $\frac{1}{2}$ von T' , welches nur mit σ_i multipliziert wird. Des Weiteren definieren wir

$$T = 2T', \quad U_1 = 2U'_1, \quad U_2 = 2U'_2 \quad \text{und} \quad V = 2V'$$

und setzen in (3.22) $c = \frac{1}{2}$. Dann erhalten wir

$$p_i(\sigma_i | \sigma_{i-1}, \sigma_{i+1}) = \frac{1}{2} (1 + \sigma_i (T + \sigma_{i+1}U_1 + \sigma_{i-1}U_2 + \sigma_{i-1}\sigma_{i+1}V)).$$

Allgemein kann man schreiben

$$p_i(\sigma_{n,i} | \sigma_n - 1) = \frac{1}{2} (1 + \sigma_{n,i} h_i(\sigma_{n-1})),$$

wobei $|h_i(\sigma_{n-1})| \leq 1$. Im Falle der oben (für die Gibbs-Maße) geforderten echt positiven Übergangswahrscheinlichkeiten gilt $|h_i(\sigma_{n-1})| < 1$. In h_i sind die Regeln des PCA enthalten und h_i ist von endlicher Reichweite.

Nun soll die Übergangswahrscheinlichkeit im Sinne der ESM-Welt mit Hilfe der Exponentialfunktion formuliert werden

$$p_i(\sigma_{n,i} | \sigma_{n-1,i-1}, \sigma_{n-1,i+1}) = p_i(\sigma_i | \sigma_{i-1}, \sigma_{i+1}) = \exp(-\Phi_x(\sigma_x, \sigma_{n-1}))$$

mit $x = (n, i) \in \mathbb{Z}^{1+1}$. Es soll gelten

$$p_i(\sigma_i | \sigma_{i-1}, \sigma_{i+1})$$

$$= \lambda^{-1} \exp(\sigma_{i-1}H_1 + \sigma_{i+1}H_2 + \sigma_i H_3 + \sigma_{i-1}\sigma_{i+1}K_1 + \sigma_i\sigma_{i+1}K_2 + \sigma_{i-1}\sigma_i K_3 + \sigma_{i-1}\sigma_{i+1}\sigma_i K_0).$$

Wieder müssen die Größen H_l , $l = 1, 2, 3$ und K_j , $j = 0, 1, 2, 3$ mit Hilfe von x, y_n , $n = 1, 2$ und z ausgedrückt werden. Mit (3.20) und (3.21) erhält man

$$\begin{aligned} z &= \lambda^{-1} \exp(H_1 + H_2 + H_3 + K_1 + K_2 + K_3 + K_0), \\ x &= \lambda^{-1} \exp(-H_1 - H_2 + H_3 + K_1 - K_2 - K_3 + K_0), \\ y_1 &= \lambda^{-1} \exp(-H_1 + H_2 + H_3 - K_1 + K_2 - K_3 - K_0), \\ y_2 &= \lambda^{-1} \exp(H_1 - H_2 + H_3 - K_1 - K_2 + K_3 - K_0), \\ 1 - z &= \lambda^{-1} \exp(H_1 + H_2 - H_3 + K_1 - K_2 - K_3 - K_0), \\ 1 - x &= \lambda^{-1} \exp(-H_1 - H_2 - H_3 + K_1 + K_2 + K_3 - K_0), \\ 1 - y_1 &= \lambda^{-1} \exp(-H_1 + H_2 - H_3 - K_1 - K_2 + K_3 + K_0), \\ 1 - y_2 &= \lambda^{-1} \exp(H_1 - H_2 - H_3 - K_1 + K_2 - K_3 + K_0). \end{aligned} \tag{3.23}$$

Da alle Größen H_l , $l = 1, 2, 3$ und K_j , $j = 0, 1, 2, 3$ je viermal mit negativem und viermal mit positivem Vorzeichen in den Gleichungen (3.23) auftauchen, erhalten wir die Normierungskonstante λ aus dem Produkt

$$\lambda^{-8} = x(1-x)z(1-z) \prod_i y_i(1-y_i).$$

Die Vorzeichenverteilung können wir auch zur Berechnung der übrigen Konstanten nutzen, so ergibt sich z. B.

$$\exp(8H_1) = \frac{zy_2(1-z)(1-y_2)}{xy_1(1-x)(1-y)}$$

und analog alle anderen Größen.

3.8. Eine Klasse von zellularen Automaten

Geben wir ein Maß auf einer Menge $\{0, 1\}^{\mathbb{Z}^d}$ vor, so gibt es keinen kanonischen Weg, einen PCA zu konstruieren, für welchen das Maß stationär ist. Dawson konnte 1975 in [10] noch spezieller zeigen, dass es Gibbs-Maße μ auf $\{0, 1\}^{\mathbb{Z}^2}$, definiert durch ein Nächster-Nachbar-Potential, gibt, die für keinen PCA stationäre reversible Maße sind. Deshalb komme ich nun von der anderen Seite und stelle eine Klasse von zellularen Automaten vor, die u. a. Gibbs-Maße als stationäre Maße besitzt, siehe [41], [56] und [74]. Sei E eine endliche Menge, z. B. $E = \{-1, +1\}$, und S eine endliche oder abzählbare Menge, wie z. B. $S = \mathbb{Z}^d$. Der Zustandsraum sei $\Omega = E^S$ und enthalte $\sigma = (\sigma_i)_{i \in S}$ mit $\sigma_i \in E$. Mit \mathfrak{F} bezeichne ich die durch die Zylinderereignisse erzeugte σ -Algebra auf Ω . Es sei $N(i) := \{j \mid |i - j| < b\} \subset S$ die endliche Nachbarschaft von $i \in S$ und $\sigma_\Lambda \in E^\Lambda$ die Restriktion von σ auf $\Lambda \in \mathcal{S}$. Mit $\sigma_{i,n}$ bezeichnen ich den Zustand an der Stelle $i \in S$ zur Zeit $n \in \mathbb{N}$. Durch $p_i(\sigma_{i,n+1} | \omega_{N(i),n})$ wird die Wahrscheinlichkeit für σ_i zur Zeit $n+1$ unter der Bedingung $\omega_{N(i)}$ zur Zeit n gegeben. Die Übergangswahrscheinlichkeiten p_i beschreiben mit

$$P_\Lambda(\sigma_{n+1,\Lambda} | \sigma_n) = \prod_{i \in \Lambda} p_i(\sigma_{n+1,i} | \sigma_{n,N(i)})$$

eine synchrone Markov-Kette. Im Allgemeinen interessiert uns der Markov-Prozeß bzw. der zellulare Automat $(\dot{\Omega} = \Omega^{\mathbb{N}}, \dot{\mathfrak{F}} = \mathfrak{F}^{\mathbb{N}}, (\sigma_n)_{n \in \mathbb{N}})$, kurz σ_n , mit $\sigma_n : \dot{\Omega} \rightarrow \Omega$, den ich mit Hilfe einer Familie von Wahrscheinlichkeitskernen P definiere. Die Kerne P werden durch lokale Übergangswahrscheinlichkeiten p_i angegeben:

$$P(d\sigma_{n+1} | \sigma_n) = \prod_{i \in S} p_i(d\sigma_{n+1,i} | \sigma_{n,N(i)}). \quad (3.24)$$

Dies bedeutet, dass der Wahrscheinlichkeitskern $P(d\sigma | \omega)$ ein Produktmaß ist mit $P_\Lambda(d\sigma | \omega) = \otimes_{i \in \Lambda} P_i(d\sigma | \omega)$ und alle σ_i , $i \in S$ gleichzeitig und unabhängig, also parallel, verändert werden, im Gegensatz zu Abschnitt 3.5. Ein Maß $\mu \in \mathcal{P}(\Omega, \mathfrak{F})$ ist ein reversibles Maß zum durch P definierten Markov-Prozeß, wenn

$$P(d\sigma | \omega) \mu(d\omega) = P(d\omega | \sigma) \mu(d\sigma) \quad \forall \sigma, \omega \in \Omega$$

gilt. Jedes reversible Maß ist ein stationäres Maß.

Für unsere Betrachtungen, die bereits im letzten Abschnitt angedeutet wurden, ist es wichtig, zwei aufeinander folgende Zeitpunkte zu betrachten. Deshalb erweitern wir S zu $\bar{S} = \{(i, r) \mid i \in S, r \in \{0, 1\}\}$. $\bar{S} = S_0 \cup S_1$ mit $S_r = \{(i, r) \mid i \in S\}$, $r \in \{0, 1\}$ stellt also die Verdopplung von S dar.

Bemerkung 3.8.1. Betrachten wir nun ein stationäres Maß μ des Markov-Prozesses mit Zustandsraum Ω , so erhalten wir ein Maß $\bar{\mu}$ auf $\bar{\Omega} = E^{\bar{S}}$. Seien $\Lambda, \Gamma \in \mathcal{S}$, wobei \mathcal{S} wieder die Menge der endlichen Teilmengen von S bezeichnet, mit $\Lambda \subset S_1$, $\bar{N}(\Lambda) \subset \Gamma \subset S_0$, dann gilt für $\sigma, \omega \in \Omega$

$$P_\Lambda(d\sigma_{n+1,\Lambda} | \omega_{n,\Gamma}) = \bar{\mu}(d\sigma_\Lambda | \omega_\Gamma), \quad P_\Gamma(d\omega_{n+1,\Gamma}) = \bar{\mu}(d\omega_\Gamma).$$

Dabei bezeichnet $\overline{N}(\Lambda) = N(\Lambda) = \cup_{i \in \Lambda} N(i)$ die endliche Nachbarschaft von Λ . Die Projektion von $\overline{\mu}$ auf Ω liefert μ . Das stationäre Maß μ ist genau dann reversibel, wenn $\overline{\mu}$ invariant unter $(i, r) \rightarrow (i, 1 - r)$, $i \in S, r \in \{0, 1\}$ ist, $\overline{\mu}$ also symmetrisch ist.

Nun soll für unsere Zwecke ein Paarpotential und die symmetrische Verdopplung des Potentials definiert werden.

Definition 3.8.1. Ein Paarpotential Φ ist ein Potential mit $|A| \leq 2$

$$\Phi_A : \begin{cases} E \rightarrow \mathbb{R} & \text{wenn } A = \{i\}, i \in S \\ E \times E \rightarrow \mathbb{R} & \text{wenn } A = \{i, j\}, i, j \in S \end{cases}$$

mit $\Phi_{\{i,j\}}(x, y) = \Phi_{\{j,i\}}(y, x)$ für $i, j \in S$ und $x, y \in E$.

Bemerkung 3.8.2. Man vergleiche mit **Definition 2.3.2**.

Haben wir ein Potential Φ auf S , so erhalten wir durch folgende symmetrische Verdopplung ein Potential $\overline{\overline{\Phi}}$ auf \overline{S} :

$$\begin{aligned} \overline{\overline{\Phi}}_{(i,0)}(x) &= \overline{\overline{\Phi}}_{(i,1)}(x) = \Phi_i(x) \\ \overline{\overline{\Phi}}_{(i,0)(j,1)}(x, y) &= \overline{\overline{\Phi}}_{(i,1)(j,0)}(x, y) = \Phi_{i,j}(x, y). \end{aligned}$$

Analog zu **Definition 2.13** können wir nun folgende Definition vornehmen, siehe [41]:

Definition 3.8.2. Ein Kern P heißt Gibbsscher Kern mit Paarpotential Φ , wenn die bedingten Wahrscheinlichkeiten p_i folgende Form haben

$$p_i(\sigma_i | \eta_{N(i)}) = \frac{\exp\left(-\Phi_i(\sigma_i) - \sum_{j \in N(i)} \Phi_{i,j}(\sigma_i, \eta_j)\right)}{\sum_{\sigma_i \in E} \exp\left(-\Phi_i(\sigma_i) - \sum_{j \in N(i)} \Phi_{i,j}(\sigma_i, \eta_j)\right)},$$

mit $i \in S$, $\sigma_i \in E$ und $\eta_{N(i)} \in E^{N(i)}$.

Satz 3.8.1. Der Kern P des Prozesses $(\sigma_n)_{n \in \mathbb{N}}$ sei strikt positiv. Dann existiert genau dann mindestens ein reversibles Maß unter den stationären Maßen des Prozesses, wenn P ein Gibbsscher Kern mit Paarpotential Φ ist.

Beweis. Haben wir ein reversibles Maß μ des Prozesses, der durch Übergangswahrscheinlichkeiten p_i (siehe (3.24)) definiert wird, die nur von der endlichen Nachbarschaft von i zur Zeit $n - 1$ abhängen, so ist das stochastische Feld mit dem entsprechenden symmetrischen Maß $\overline{\mu}$ auf $\overline{\Omega}$ ein Markov- und damit ein Gibbs-Feld mit symmetrischem $\overline{\overline{\Phi}}$. Also bezeichnet das symmetrische Potential wegen der Eindeutigkeit von $\overline{\overline{\Phi}}$ die symmetrische Verdopplung eines Paarpotentials Φ . Durch P ist ein Gibbsscher Kern mit dem Paarpotential Φ gegeben.

Haben wir einen Gibbsschen Kern und damit ein Paarpotential Φ , so erhalten wir durch eine symmetrische Verdopplung $\overline{\overline{\Phi}}$ ein Gibbs-Feld zu dem Potential $\overline{\overline{\Phi}}$. Mit dem symmetrischen Gibbs-Maß auf $\overline{\Omega}$ erhalten wir zu einem Markov-Prozeß mit Gibbsschem Kern ein reversibles Maß auf Ω , vgl. [41]. \square

3. Markov-Prozeß und zellulare Automaten

Bemerkung 3.8.3. Man beachte Bemerkung 2.3.3 und Kapitel 4.2 für diesen Beweis.

Der nächste Satz ermöglicht es mir, zu zeigen, dass mit den Voraussetzungen des **Satzes (3.8.1)** jedes reversible Maß eines Prozesses ein Gibbs-Maß zu einem bestimmten Potential ist, vgl. [41].

Satz 3.8.2. *Sei $\bar{\mu}$ ein strikt positives Gibbs-Maß auf dem doppelten Graphen $(S_0 \cup S_1)$ mit Potential $\bar{\Phi}$ der Reichweite einer endlichen Nachbarschaft $\bar{N}(i)$, dann gilt*

- für $\Lambda, \Gamma \in \mathcal{S}$, $\Lambda \subset S_1$, $\bar{N}(i) \subset \Gamma \subset S_0$, $\sigma_\Lambda \in E^\Lambda$, $\eta_\Gamma \in E^\Gamma$

$$\bar{\mu}(\sigma_\Lambda | \eta_\Gamma) = \prod_{i \in \Lambda} \frac{\exp\left(-H_i^{\bar{\Phi}}(\sigma_i, \eta_{\bar{N}(i)})\right)}{\sum_{j \in E} \exp\left(-H_i^{\bar{\Phi}}(j, \eta_{\bar{N}(i)})\right)}$$

- die Projektion μ von $\bar{\mu}$ auf $\Omega = E^{S_0}$ ist ein Gibbs-Maß mit dem Potential $\hat{\Phi}$, definiert durch

$$\begin{aligned} \hat{\Phi}_i(\sigma_i) &= \bar{\Phi}_i(\sigma_i), \quad i \in S \\ \hat{\Phi}_{\bar{N}(i)}(\sigma_{\bar{N}(i)}) &= -\ln \sum_{\sigma_i \in E} \exp\left(-H_i^{\bar{\Phi}}(\sigma_i, \eta_{\bar{N}(i)})\right). \end{aligned}$$

Bemerkung 3.8.4. $H_i^{\bar{\Phi}}$ bezeichnet analog zu **Definition 2.3.2** wieder die Hamilton-Funktion. Für ein Paarpotential gilt

$$H_i^{\bar{\Phi}}(\sigma_i, \eta_{\bar{N}(i)}) = \Phi_i(\sigma_i) + \sum_{j \in \bar{N}(i)} \Phi_{i,j}(\sigma_i, \eta_j).$$

Natürlich läßt sich auch diese Schreibweise auf Potentiale verallgemeinern. Das Potential Φ müßte also kein Paarpotential sein!

$\bar{N}(i)$ und $N(i)$ sind identisch, deshalb werde ich im folgenden $N(i)$ nutzen.

Beweis. 1. Punkt: klar nach Definition von Gibbs-Maßen!

2. Punkt: [41] □

Satz 3.8.3. *Wenn ein Prozeß $(\dot{\Omega} = \Omega^\mathbb{N}, \dot{\mathfrak{F}} = \mathfrak{F}^\mathbb{N}, \mu, (\sigma_n)_{n \in \mathbb{N}})$ reversibel mit Gibbsschem Kern und Potential Φ ist, dann ist das reversible Maß μ ein Gibbs-Maß mit Potential $\hat{\Phi}$.*

Beweis. Haben wir ein reversibles Maß μ des Markov-Prozesses mit Gibbsschem Kern, so ist das entsprechende Maß $\bar{\mu}$ auf $\bar{\Omega}$ ein Gibbs-Maß. Die Projektion Ω ist unser reversibles Maß μ und nach Satz 3.8.2 ein Gibbs-Maß mit Potential $\hat{\Phi}$. □

Zusammenfassend können wir sagen, dass jedes reversible Maß des Markov-Prozesses mit Gibbsschem Kern ein Gibbs-Maß bezüglich $\hat{\Phi}$ ist. Weiterhin kann gezeigt werden, dass jedes Gibbs-Maß μ bezüglich $\hat{\Phi}$ reversibel oder 2-periodisch ist, $\mu(\mu P) = \mu$, siehe [41], [56] und dass zumindest ein Gibbs-Maß bezüglich $\hat{\Phi}$ ein reversibles Maß des Markov-Prozesses ist.

Teil II.

Markov-Felder und Markov-Ketten

4. Zusammenhänge von Markov-Feldern, Markov-Ketten und Gibbs-Maßen

In diesem Kapitel zeige ich, wie man Markov-Ketten als Gibbs-Maße identifizieren kann. Dafür müssen zum einen Grundlagen der Markov-Ketten, wie z. B. der Ergodensatz, zur Verfügung gestellt werden. Und zum anderen benötigen wir einige Ergänzungen zu stochastischen Feldern.

4.1. Ergodensatz für Markov-Ketten

Es sei $P = (P(x, y))_{x, y \in E}$ eine positive Markov-Matrix (=stochastische Matrix), d. h. die Einträge der Matrix P^k sind für ein $k \in \mathbb{N}$ positiv (>0) und die Spaltensumme der Einträge der Matrix P ist 1, mit

$$P \cdot p(i) = p(i+1). \quad (4.1)$$

Der Zustandsraum E sei endlich mit $|E| = n$. Die Wahrscheinlichkeitsverteilung $p(i) = (p_1(i), \dots, p_n(i)) \in]0, 1[^E$ gibt die Wahrscheinlichkeit für die Zustände zur Zeit i bzw. an der Stelle $i \in S = \mathbb{Z}$ an. S ist die Parametermenge der Markov-Kette. Bisher haben wir die Parametermenge immer als Zeit interpretiert. $P(x, y)$ gibt die Übergangswahrscheinlichkeit von Zustand x zu Zustand y in einem Zeitschritt an. Mit Blick auf den Zusammenhang zwischen Markov-Ketten und Gibbs-Maßen interpretieren wir die Parametermenge nun als eindimensionales Gitter. $P(x, y)$ gibt dann die Wahrscheinlichkeit von Zustand y an der Stelle $i+1$ an, wenn an der Stelle i Zustand x ist. Positive Markov-Matrizen heißen regulär oder ergodisch. Wir wollen nun den Ergodensatz formulieren, siehe [26] und [17].

Satz 4.1.1. *Für eine reguläre bzw. ergodische Markov-Matrix P existiert eine eindeutige Wahrscheinlichkeitsverteilung α_P mit*

$$P\alpha_P = \alpha_P \quad \text{und} \quad \alpha_P(x) > 0, \quad x \in E.$$

Die Verteilung α_P wird stationäre Verteilung genannt. Für jede Wahrscheinlichkeitsverteilung $p(0)$ gilt

$$\lim_{i \rightarrow \infty} p(i) = \lim_{i \rightarrow \infty} P^i p(0) = \alpha_P.$$

Bemerkung 4.1.1.

4. Zusammenhänge von Markov-Feldern, Markov-Ketten und Gibbs-Maßen

- Eine Markov-Kette ist irreduzibel, wenn jeder Zustand von jedem Zustand aus erreichbar ist, wobei mehrere Schritte erlaubt sind. Es muss nicht ein $k \in \mathbb{N}$ existieren, so dass alle Einträge der Matrix P^k positiv sind. So ist zum Beispiel die 2-periodische Matrix

$$\begin{pmatrix} 0 & 1 & 1 \\ \frac{1}{2} & 0 & 0 \\ \frac{1}{2} & 0 & 0 \end{pmatrix}$$

irreduzibel.

- Allgemeiner kann man für aperiodische irreduzible Markov-Ketten analog zu den kontinuierlichen Markov-Ketten, siehe **Bemerkung 3.4.4**, folgendes formulieren: Hat das System $P\alpha_P = \alpha_P$ eine eindeutige Lösung α_P mit $\sum \alpha_P(x) = 1$ und $\alpha_P(x) > 0 \forall x \in E$, dann gilt $\lim_{i \rightarrow \infty} p(i) = \alpha_P$. Ist der Zustandsraum endlich, so existiert immer eine eindeutige Lösung α_P mit $\sum \alpha_P(x) = 1$ und $\alpha_P(x) > 0 \forall x \in E$, da aperiodische irreduzible Markov-Ketten einer positiven Markov-Matrix entsprechen. Betrachten wir einen abzählbaren Zustandsraum, so kann es passieren, dass $\lim_{i \rightarrow \infty} p(i) = 0$ gilt, vgl. [51].

Der Beweis des Ergodensatzes benötigt den Satz von Perron und Frobenius, der hier zitiert werden soll, vgl. [26].

Satz 4.1.2. *Jede positive Matrix $Q = (Q(x, y))_{x, y \in E}$ mit endlicher Menge $|E| \geq 2$ hat einen einfachen Eigenwert $q > 0$ mit den folgenden Eigenschaften.*

- Für alle Eigenwerte $z \neq q$ von Q gilt $|z| < q$.
- Es gibt zu q einen rechten Eigenvektor r mit positiven Einträgen $r(x) > 0, x \in E$.
- Sind r und r' Eigenvektoren zum Eigenwert q , dann gilt $r' = cr$ für ein $c \in \mathbb{C}$.

Man bezeichnet q als PF-Eigenwert von Q .

Bemerkung 4.1.2. Der PF-Eigenwert einer ergodischen Markov-Matrix ist 1. Sei r der rechte Eigenvektor zum PF-Eigenwert, dann ergibt sich

$$\alpha_P = \frac{1}{\sum_{x \in E} r(x)} r.$$

Jetzt kommen wir zum Beweis des **Satzes 4.1.1**.

Beweis. Im Beweis gehe ich ohne Einschränkung davon aus, dass P bereits eine positive Matrix ist, vgl. Anhang von [38].

1.) Der PF-Eigenwert von P ist 1.

Es sei q der PF-Eigenwert von P und r der dazugehörige Eigenvektor. Dann gilt

$$\sum_{x \in E} P(x, y) r(x) = q r(y), \quad y \in E.$$

Wird über y summiert folgt

$$1 = q \cdot 1.$$

Und damit ist der PF-Eigenwert 1.

2.) Die Wahrscheinlichkeitsverteilung α_P ist eindeutig.

Der Eigenvektor r zum Eigenwert 1 kann nach dem Satz von PF positiv gewählt werden.

Die Eigenvektoren zu q unterscheiden sich nach PF nur durch eine Konstante.

3.) Für jede Wahrscheinlichkeitsverteilung $p(0)$ gilt

$$\lim_{i \rightarrow \infty} p(i) = \lim_{i \rightarrow \infty} P^i p(0) = \alpha_P.$$

Es seien $\lambda_1, \dots, \lambda_n$ die Eigenwerte der Matrix P mit $\lambda_1 = 1$ und $|\lambda_j| < 1$, $j = 2, \dots, n$.

Damit hat P die Jordan-Form

$$J = \begin{pmatrix} 1 & 0 \\ 0 & J_\star \end{pmatrix},$$

wobei J_\star die Eigenwerte $\lambda_2, \dots, \lambda_n$ besitzt. Damit gilt

$$\lim_{i \rightarrow \infty} J^i = \text{diag}(1, 0, \dots, 0).$$

Sei nun $P = QJ^nQ^{-1}$, wobei die erste Spalte von Q der Eigenvektor r von P zum Eigenwert 1 ist, dann erhalten wir

$$\begin{aligned} \lim_{i \rightarrow \infty} p(i) &= \lim_{i \rightarrow \infty} P^i p(0) = \lim_{i \rightarrow \infty} QJ^iQ^{-1}p(0) = \lim_{i \rightarrow \infty} (QJ^i)Q^{-1}p(0) \\ &= (r, 0, \dots, 0) \cdot (a, \star, \dots, \star)^T = a \cdot r = \alpha_P \end{aligned}$$

Zusammen mit $\sum_{x \in E} p_x(i) = 1$ erhalten wir

$$\sum_{x \in E} ar(x) = 1$$

und damit

$$a = \frac{1}{\sum_{x \in E} r(x)}.$$

□

4.2. Stochastische Felder

Ich habe bereits herausgestellt, dass Familien von Wahrscheinlichkeitskernen sowohl für die Definition der Gibbs-Maße als auch für die Definition des Markov-Prozesses eine bedeutende Rolle spielen. Ein wesentlicher Unterschied ist jedoch, dass die Kerne im Fall des Markov-Prozesses durch die Zeit und im Fall der Gibbs-Maße räumlich indiziert werden. Ist das Gitter nur eindimensional, z. B. $S = \mathbb{Z}$ kann aber unter bestimmten Voraussetzungen gezeigt werden, dass die Gibbssche Spezifizierung eine Markovsche Halbgruppe ist und damit einen Markov-Prozeß definiert. Die unterschiedlichen Indizierungen durch die Zeit bzw. den Raum möchte ich durch folgende Wiederholung untermauern.

4. Zusammenhänge von Markov-Feldern, Markov-Ketten und Gibbs-Maßen

Definition 4.2.1. Das Quadrupel $(\Omega, \mathfrak{F}, P, (\pi_i)_{i \in S})$ heißt stochastisches Feld, wobei $(\pi_i)_{i \in S}$ eine Familie von Zufallsvariablen mit dem gemeinsamen Zustandsraum (E, \mathfrak{E}) ist. S ist die Parametermenge und $(\Omega, \mathfrak{F}, P)$ ein Wahrscheinlichkeitsraum.

Wir setzen $\Omega = E^S$, wobei $E = \{0, 1\}$ und S eine abzählbar unendliche Menge ist, z. B. \mathbb{Z}^d , des Weiteren sei π_i die Projektion

$$\begin{aligned} \pi_i : \Omega &\rightarrow E \\ \omega &\mapsto \omega_i \end{aligned}$$

mit $\omega = (\omega_i)_{i \in S} \in \Omega$. Damit können wir stochastische Felder auf unsere Gittergase beziehen. Ist eine Spezifizierung, eine Familie von Wahrscheinlichkeitskernen, und ein Startmaß für den Aussenraum gegeben, lässt sich ein Wahrscheinlichkeitsmaß auf Ω konstruieren. Die Gibbsche Spezifizierung spielt für Markov-Felder eine ähnliche Rolle wie die Markovsche Halbgruppe für den Markov-Prozess. Jedoch haben wir eine räumliche Indizierung. Die Markovsche Halbgruppe diente dazu, eine projektive Familie von W-Maßen zu erzeugen. Analog können wir mit der Gibbsschen Spezifizierung verfahren

$$P_\Omega(A) = \int_{E^{S \setminus \Lambda}} \gamma_\Lambda^\Phi(A | \omega_{S \setminus \Lambda}) P_{S \setminus \Lambda}(\omega_{S \setminus \Lambda}) \quad \forall A \in \mathfrak{F}, \omega \in \Omega, \Lambda \in \mathcal{S}. \quad (4.2)$$

\mathcal{S} spielt hier die Rolle von $\mathfrak{h}(T)$ und $P_{S \setminus \Lambda}$ ist das auf den Aussenraum eingeschränkte Gibbs-Maß. Das auf diese Weise analog zu Satz 3.1.1 zu konstruierende W-Maß auf Ω ist das nicht eindeutige Gibbs-Maß.

So wie die Markovschen Halbgruppen einen Markov-Prozeß definieren, definieren die Gibbsschen Spezifizierungen ein spezielles stochastisches Feld, das Gibbs-Feld.

Definition 4.2.2. Ein stochastisches Feld, dessen W-Maß P die Bedingung

$$P(A | \mathfrak{F}_\Lambda) = P(A | \mathfrak{F}_{\partial\Lambda}) \quad P\text{-fast sicher}$$

mit $A \in \mathfrak{F}_\Lambda$ und $\Lambda \in \mathcal{S}$ erfüllt, heißt Markov-Feld.

Bemerkung 4.2.1. Ich setze $\partial\Lambda = \{i \in S \setminus \Lambda \mid \exists j \in \Lambda \mid |i - j| = 1\}$

In [65] wird gezeigt, dass jedes Markov-Feld ein Gibbs-Feld und umgekehrt ist, wenn

- beide translationsinvariant sind,
- das Markov-Feld von erster Ordnung ist, d. h. $\partial\Lambda$ besteht nur aus den Gitterpunkten, die zu mindestens einem Gitterpunkt in Λ den Abstand 1 haben,
- das Gibbs-Feld durch ein Nächster-Nachbar-Potential mit binärer Wechselwirkung definiert ist,
- $S = \mathbb{Z}^d$,
- E aus zwei Zuständen besteht.

Sherman verallgemeinert in [64] und zeigt, dass die fünf Bedingungen vernachlässigt werden können und jedes Gibbs-Feld trotzdem ein Markov-Feld höherer Ordnung und umgekehrt ist, vgl. auch [7]. Für Potentiale mit Interaktion zwischen mehr als zwei Teilchen benötigt man Markov-Felder höherer Ordnung, d. h. $\partial\Lambda$ wird größer. Ich gehe im nächsten Abschnitt 4.3 auf die Äquivalenz für einen Spezialfall mit $S = \mathbb{Z}$ ein. In dem Spezialfall können wir nämlich zusätzlich eine Äquivalenz von Markov-Feldern und Markov-Ketten beweisen.

4.3. Homogene Markov-Spezifizierungen und Gibbssche Spezifizierungen

Ich möchte jetzt $S = \mathbb{Z}$ betrachten und dann die Gibbs-Felder als Markov-Felder bzw. Markov-Kette identifizieren, vgl. [26]. Im eindimensionalen Gitter sind die nächsten Nachbarn eines Gitterpunktes i nur die Gitterpunkte links und rechts von i . Es gilt die zweiseitige Markov-Bedingung

$$P(A|\mathfrak{T}_\Lambda) = P(A|\mathfrak{F}_{\{i-1, k+1\}}) \quad P\text{-fast sicher}$$

mit $\Lambda = \{i, i+1, \dots, k\} \in \mathbb{Z}$. Für den Markov-Prozeß gilt die einseitige Markov-Bedingung

$$P(A|\mathfrak{T}_\Lambda) = P(A|\mathfrak{F}_{\{i-1\}}) \quad P\text{-fast sicher.}$$

Die Indizierung ist noch räumlich, jedoch eindimensional, so dass es mathematisch gleich ist, ob mit der Parametermenge die Zeit oder der eindimensionale Raum bezeichnet wird. Jeder Markov-Prozeß ist also ein Markov-Feld. Die Umkehrung gilt, wenn E endlich und das Markov-Feld positiv und homogen ist. Für den Fall eines nicht-endlichen Zustandsraums verweise ich auf [26]. Als Beispiel für einen Markov-Prozess auf einem Gitter möchte ich nun einen räumlich indizierten wählen, nämlich $T = S = \mathbb{Z}$, $E = \{0, 1\}$ und $\Omega = E^S$. Da der Zustandsraum endlich und die Parametermenge abzählbar ist, handelt es sich um eine Markov-Kette. Ich werde zeigen, dass diese Markov-Kette einem Gibbs-Maß entspricht. Die Markov-Kette wird in diesem Fall auch ein Markov-Feld sein. Beginnen wir mit der Betrachtung eines Markov-Feldes.

Definition 4.3.1. Es sei γ eine Spezifizierung mit Parametermenge $S = \mathbb{Z}$ und dem Zustandsraum E . γ ist eine positive homogene Markov-Spezifizierung erster Ordnung, wenn es eine Funktion $g(\cdot, \cdot, \cdot) > 0$ auf E^3 gibt, so dass

$$\gamma_{\{i\}}(\pi_i = y | \omega) = g(\omega_{i-1}, y, \omega_{i+1}) \quad \forall i \in \mathbb{Z}, y \in E, \omega \in \Omega.$$

Nach der Definition bestimmt g nur $(\gamma_{\{i\}})_{i \in S}$. Der folgende Satz zeigt, dass dies Auswirkungen auf $(\gamma_\Lambda)_{\Lambda \in \mathcal{S}}$ hat und g als γ -bestimmende Funktion bezeichnet werden kann, siehe [26], [19], [20].

Satz 4.3.1. Es seien $\lambda \in \mathcal{M}(E, \mathfrak{E})$ und γ eine Spezifizierung mit der folgenden Eigenschaft:

4. Zusammenhänge von Markov-Feldern, Markov-Ketten und Gibbs-Maßen

Für alle $i \in S$ gibt es eine messbare Funktion $\rho_{\{i\}} > 0$ auf Ω , so dass $\gamma_{\{i\}} = \rho_{\{i\}} \lambda_{\{i\}}$. Dann gibt es eine positive λ -Modifikation ρ mit $\gamma = \rho \lambda$. Die Spezifizierung γ ist eindeutig bestimmt durch $(\rho_{\{i\}})_{i \in S}$, λ und

$$\begin{aligned} \mathcal{G}(\gamma) &= \{ \mu \in \mathcal{P}(\Omega, \mathfrak{F}) : \mu \gamma_{\{i\}} = \mu \quad \forall i \in S \} \\ &= \{ \mu \in \mathcal{P}(\Omega, \mathfrak{F}) : \mu = \rho_{\{i\}}(\mu \lambda_{\{i\}}) \quad \forall i \in S \}. \end{aligned}$$

Beweis. Unsere Voraussetzungen sind,

a) dass γ eine Spezifizierung mit der Eigenschaft $\gamma_{\{i\}} = \rho_{\{i\}} \lambda_{\{i\}} \quad \forall i \in S$ ist, wobei $\rho_{\{i\}} > 0$ eine messbare Funktion ist.

b) dass $\lambda \in \mathcal{M}(E, \mathfrak{E})$ und mit **Definition 2.3.1** können wir sogar $\lambda \in \mathcal{P}(E, \mathfrak{E})$ voraussetzen.

Jetzt definieren wir für jedes $\Lambda \in \mathcal{S}$ mit $(\rho_{\{i\}})_{i \in \Lambda}$ und $\lambda \in \mathcal{P}(E, \mathfrak{E})$ eine messbare Funktion $\rho_\Lambda > 0$, so dass

$$\mu = \rho_\Lambda(\mu \lambda_\Lambda)$$

für alle $\mu \in \mathcal{P}(\Omega, \mathfrak{F})$, die $\mu \gamma_{\{i\}} = \mu \quad \forall i \in \Lambda$ erfüllen, gilt.

Nach Voraussetzung ist γ eine Spezifizierung und erfüllt damit die Konsistenzbedingung

$$\gamma_\Lambda(\cdot | \omega) \gamma_{\{i\}} = \gamma_\Lambda(\cdot | \omega) \quad \forall i \in \Lambda, \quad \forall \omega \in \Omega.$$

Damit haben wir aufgrund der Konstruktion von ρ_Λ

$$\gamma_\Lambda(\cdot | \omega) = \rho_\Lambda(\gamma_\Lambda(\cdot | \omega) \lambda_\Lambda) = \rho_\Lambda \lambda_\Lambda(\cdot | \omega) \quad \forall \omega \in \Omega.$$

Also gilt $\gamma_\Lambda = \rho_\Lambda \lambda_\Lambda$ und wir haben gezeigt, dass $(\rho_\Lambda)_{\Lambda \in \mathcal{S}}$ eine λ -Modifikation mit $\gamma = \rho \lambda$ ist und γ durch $(\rho_{\{i\}})_{i \in S}$ sowie λ eindeutig bestimmt wird. Außerdem gilt für $\mu \in \mathcal{P}(\Omega, \mathfrak{F})$ mit $\mu \gamma_{\{i\}} = \mu$ für alle $i \in S$

$$\mu = \rho_\Lambda(\mu \lambda_\Lambda) = \mu \gamma_\Lambda \quad \forall \Lambda \in \mathcal{S}$$

und damit $\mu \in \mathcal{G}(\gamma)$.

Abschließend müssen wir noch die oben angekündigte Definition von ρ_Λ für jedes $\Lambda \in \mathcal{S}$ liefern. Dies geschieht per Induktion über $|\Lambda|$. Der Fall $|\Lambda| = 1$ ist klar. Sei $\Lambda \in \mathcal{S}$ die Vereinigung von disjunkten Mengen $\Lambda_1, \Lambda_2 \in \mathcal{S}$ und seien ρ_{Λ_1} und ρ_{Λ_2} bereits definiert, so setzen wir

$$\rho_\Lambda = \begin{cases} \frac{\rho_{\Lambda_1}}{\lambda_{\Lambda_1}(\rho_{\Lambda_1} \rho_{\Lambda_2}^{-1})} & \text{wenn } \lambda_{\Lambda_1}(\rho_{\Lambda_1} \rho_{\Lambda_2}^{-1}) < \infty, \\ 1 & \text{sonst.} \end{cases}$$

Dann gilt $0 < \rho_\Lambda < \infty$ und ρ_Λ ist messbar. Nun sei $\mu \in \mathcal{P}(\Omega, \mathfrak{F})$ mit $\mu \gamma_{\{i\}} = \mu \quad \forall i \in \Lambda$, dann ist $\mu = \rho_\Lambda(\mu \lambda_\Lambda)$ zu zeigen. Nach Voraussetzung gilt $\mu = \rho_{\Lambda_k}(\mu \lambda_{\Lambda_k})$, μ ist also konsistent mit $\rho_{\Lambda_k} \lambda_{\Lambda_k}$, für $k = 1, 2$. Sei $f \geq 0$ eine beliebige messbare Funktion, dann gilt

$$\mu((\rho_\Lambda \lambda_\Lambda)(f)) = \mu(\lambda_{\Lambda_2}((\rho_\Lambda \lambda_{\Lambda_1})(f))).$$

4.3. Homogene Markov-Spezifizierungen und Gibbssche Spezifizierungen

Jetzt nutzen wir die induktive Definition von ρ_Λ und erhalten für die rechte Seite

$$\mu \left[\lambda_{\Lambda_2} \left(\left(\frac{\rho_{\Lambda_1}}{\lambda_{\Lambda_1}(\rho_{\Lambda_1}\rho_{\Lambda_2}^{-1})} \lambda_{\Lambda_1} \right) (f) \right) \right].$$

Nach Voraussetzung ist μ konsistent mit $\rho_{\Lambda_k} \lambda_{\Lambda_k}$ für $k = 1, 2$ und damit folgt

$$\mu \left[(\rho_{\Lambda_1} \lambda_{\Lambda_1}) \left(\rho_{\Lambda_2}^{-1} \left(\frac{\rho_{\Lambda_1}}{\lambda_{\Lambda_1}(\rho_{\Lambda_1}\rho_{\Lambda_2}^{-1})} \lambda_{\Lambda_1} \right) (f) \right) \right].$$

Wir erhalten

$$\mu((\rho_\Lambda \lambda_\Lambda)(f)) = \mu \left[(\rho_{\Lambda_1} \lambda_{\Lambda_1}) \left(\frac{f}{\lambda_{\Lambda_1}(\rho_{\Lambda_1}\rho_{\Lambda_2}^{-1})} \lambda_{\Lambda_1}(\rho_{\Lambda_1}\rho_{\Lambda_2}^{-1}) \right) \right] = \mu((\rho_{\Lambda_1} \lambda_{\Lambda_1})(f)).$$

Damit gilt $\mu = \rho_\Lambda(\mu \lambda_\Lambda)$, da μ konsistent mit $\rho_{\Lambda_1} \lambda_{\Lambda_1}$, und es folgt $\lambda_{\Lambda_1}(\rho_{\Lambda_1}\rho_{\Lambda_2}^{-1}) < \infty$. \square

Bemerkung 4.3.1.

- Es gilt

$$\mu(f) = \int f d\mu,$$

$(f\mu)$ ist das Maß mit der Radon-Nikodym Dichte f bezüglich μ und

$$(\lambda_{\Lambda_1}(\lambda_{\Lambda_2}(f(\omega)))) = \int \lambda^{\Lambda_1}(d\zeta) \int \lambda^{\Lambda_2}(d\eta) f(\eta\zeta\omega_{S\setminus\Lambda}).$$

- Für die letzte Umformung benötigen wir die Identität

$$(\rho_{\Lambda_1} \lambda_{\Lambda_1})(\rho_{\Lambda_2}^{-1}(\rho_{\Lambda_1} \lambda_{\Lambda_1})(f)) = (\rho_{\Lambda_1} \lambda_{\Lambda_1})(f(\rho_{\Lambda_1} \lambda_{\Lambda_1})(\rho_{\Lambda_2}^{-1})).$$

- Ich möchte noch eine Bemerkung zur Konsistenz von μ mit $\rho_{\Lambda_k} \lambda_{\Lambda_k}$ für $k = 1, 2$ machen. Dazu definiere ich $\mu_{\Lambda_1} := \rho_{\Lambda_1} \lambda_{\Lambda_1}$ und $\mu_{\Lambda_2} := \rho_{\Lambda_2} \lambda_{\Lambda_2}$. Damit bedeutet die Konsistenz

$$\mu = \mu_{S\setminus\Lambda_1} \mu_{\Lambda_1} = \mu_{S\setminus\Lambda_2} \mu_{\Lambda_2}.$$

Mit der Radon-Nikodym Ableitung

$$\frac{d\mu_{\Lambda_2}}{d\lambda_{\Lambda_2}} = \rho_{\Lambda_2}$$

erhalten wir

$$d\lambda_{\Lambda_2} = d\mu_{\Lambda_2} \rho_{\Lambda_2}^{-1}.$$

Jetzt ist die Ersetzung von λ_{Λ_2} im letzten Beweis offensichtlich

$$\int d\mu_{S\setminus\Lambda_2} \int d\lambda_{\Lambda_2} = \int d\mu_{S\setminus\Lambda_2} \int d\mu_{\Lambda_2} \rho_{\Lambda_2}^{-1} = \int d\mu_{S\setminus\Lambda_2} \rho_{\Lambda_2}^{-1} = \int d\mu_{S\setminus\Lambda_1} \int d\lambda_{\Lambda_1} \rho_{\Lambda_1} \rho_{\Lambda_2}^{-1}.$$

4. Zusammenhänge von Markov-Feldern, Markov-Ketten und Gibbs-Maßen

Es sei $P = (P(x, y))_{x, y \in E}$ eine positive Markov-Matrix und $P(x_0, x_1)$ sei die Übergangswahrscheinlichkeit von Zustand $x_0 \in E$ an der Stelle $i \in \mathbb{Z}$ zu $x_1 \in E$ an der Stelle $i + 1$. Mit der Matrix P und dem Vektor α_P kann ein W-Maß μ_P auf $\Omega = E^S$ definiert werden durch

$$\mu_P(\pi_i = x_0, \pi_{i+1} = x_1, \dots, \pi_{i+n} = x_n) = \alpha_P(x_0)P(x_0, x_1) \dots P(x_{n-1}, x_n). \quad (4.3)$$

Es gibt die Wahrscheinlichkeit an, dass an der Stelle i der Zustand $x_0 \in E$, an der Stelle $i + 1$ der Zustand x_1 usw. angenommen wird. Mit $\alpha_P(x_0)$ bezeichnen wir die dem bei uns zweidimensionalen Vektor α_P entnommene Wahrscheinlichkeit für den Zustand $x_0 = 0$ oder 1 . Den Zusammenhang zwischen der ein Markov-Feld definierenden Markov-Spezifizierungen und dem Markov-Prozeß bzw. der Markov-Kette liefert der folgende Satz, vgl. [66], [26]. Für die Formulierung des Satzes benötige ich zuvor die Definition des Randes $\partial\Lambda$ einer endlichen Teilmenge von \mathbb{Z}

$$\partial\Lambda = \{i \in \mathbb{Z} \setminus \Lambda : |i - j| = 1 \text{ für irgendwelche } j \in \Lambda\}.$$

Satz 4.3.2. *Die Relation $\mathcal{G}(\gamma) = \{\mu_P\}$ etabliert eine eins-zu-eins Korrespondenz $\gamma \leftrightarrow P$ zwischen der Menge aller positiven homogenen Markov-Spezifizierungen und der Menge aller positiven Markov-Matrizen. Für eine gegebene Matrix P kann das zugehörige γ folgendermaßen bestimmt werden*

$$\gamma_\Lambda(\pi_\Lambda = \zeta \mid \omega) = \mu_P(\pi_\Lambda = \zeta \mid \pi_{\partial\Lambda} = \omega_{\partial\Lambda}) \quad \Lambda \in \mathcal{S}, \omega \in \Omega, \zeta \in E^\Lambda. \quad (4.4)$$

Die Matrix P kann umgekehrt mit der γ definierenden Funktion g ausgedrückt werden:

$$P(x, y) = \frac{Q(x, y)r(y)}{qr(x)} \quad x, y \in E. \quad (4.5)$$

Dabei ist $Q(x, y) = \frac{g(a, x, y)}{g(a, a, y)}$ für ein beliebiges aber festes $a \in E$, q ist der größte positive Eigenwert von $Q = (Q(x, y))_{x, y \in E}$ und $\mathbf{r} \in]0, \infty[^E$ ein dazugehöriger rechter Eigenvektor.

Bemerkung 4.3.2. Da wir uns auf den eindimensionalen Fall $S = \mathbb{Z}$ beschränken, können alle $\Lambda \in \mathcal{S}$ als Vereinigungen dargestellt werden:

$$\Lambda = \bigcup_{k=1}^n \{i_k + 1, \dots, i_k + n_k\}$$

mit $n \geq 1$, $i_k \in \mathbb{Z}$, $n_k \in \mathbb{N}$, so dass die Mengen $\{i_k + 1, \dots, i_k + n_k\}$ paarweise disjunkt sind. Mit (4.1) erhalten wir

$$P^{n_k+1} \cdot p(i_k) = p(i_k + n_k + 1).$$

$P^{n_k+1}(\omega_{i_k}, \omega_{i_k+n_k+1})$ soll der Eintrag der Matrix P^{n_k+1} sein, der die Übergangswahrscheinlichkeit von Zustand $\omega_{i_k} \in E$ an der Stelle i_k nach $\omega_{i_k+n_k+1} \in E$ an der Stelle

4.3. Homogene Markov-Spezifizierungen und Gibbssche Spezifizierungen

$i_k + n_k + 1$ angibt. Mit (4.3) können wir die homogene positive Markov-Spezifizierung γ_Λ zu gegebener Markov-Matrix P wie folgt angeben

$$\begin{aligned} & \mu_P(\pi_\Lambda = \zeta \mid \pi_{\partial\Lambda} = \omega_{\partial\Lambda}) \\ &= \prod_{k=1}^n \frac{P(\omega_{i_k}, \zeta_{i_k+1})P(\zeta_{i_k+1}, \zeta_{i_k+2}) \dots P(\zeta_{i_k+n_k}, \omega_{i_k+n_k+1})}{P^{n_k+1}(\omega_{i_k}, \omega_{i_k+n_k+1})}. \end{aligned} \quad (4.6)$$

Beweis. Im Folgenden bezeichne ich mit Π die Menge aller Markov-Matrizen mit nicht verschwindenden Einträgen und einer Zeilensumme von 1, d. h. wir transponieren die bisher betrachteten Matrizen P und multiplizieren $p(i)$ von links an die Matrix. Mit Γ bezeichnen wir die Menge aller positiven homogenen Markov-Spezifizierungen.

Der Beweis erfolgt nun in mehreren Schritten.

1.) Es sei zunächst $P \in \Pi$ gegeben und γ wird durch (4.4) definiert. Zu zeigen ist, dass dann $\gamma \in \Gamma$ und $\mu_P \in \mathcal{G}(\gamma)$. Wähle dafür $\Lambda \in \mathcal{S}$ fest. Dann gilt für alle $\zeta \in E^\Lambda$, $\omega \in \Omega$ und $\Delta \in \mathcal{S}$ mit $\Lambda \cup \partial\Lambda \subset \Delta$

$$\mu_P(\sigma_\Lambda = \zeta \mid \sigma_{\Delta \setminus \Lambda} = \omega_{\Delta \setminus \Lambda}) = \mu_P(\sigma_\Lambda = \zeta \mid \sigma_{\partial\Lambda} = \omega_{\partial\Lambda}). \quad (4.7)$$

Damit gilt für alle $\zeta \in E^\Lambda$ und $A \in \mathcal{T}_\Lambda$

$$\mu_P(\{\sigma_\Lambda = \zeta\} \cap A) = \int_A d\mu_P \gamma_\Lambda(\sigma_\Lambda = \zeta \mid \cdot). \quad (4.8)$$

Also gilt $\mu_P = \mu_P \gamma_\Lambda$ und damit $\mu_P \in \mathcal{G}(\gamma)$. Nach der Definition für positive homogene Markov-Spezifizierungen müssen wir noch eine bestimmende Funktion für γ angeben. Diese ergibt sich leicht mit (4.4) zu

$$g(x, y, z) = \frac{P(x, y)P(y, z)}{P^2(x, z)}, \quad (4.9)$$

wobei $P^2(x, z)$ ein Eintrag der Matrix P^2 ist.

Es gilt also $\gamma \in \Gamma$ und (4.4) definiert eine Abbildung $b : \Pi \rightarrow \Gamma$ mit $\mu_P \in \mathcal{G}(b(P))$ für alle $P \in \Pi$.

2.) Jetzt zeigen wir für $P \in \Pi$ und $\gamma = b(P)$, dass man P durch Gleichung (4.5) mit γ erhält. Ausserdem zeigen wir, dass b injektiv ist.

Es sei $a \in E$ fest gewählt. Die γ -bestimmende Funktion g haben wir im ersten Teil definiert, so dass mit der Definition von Q für alle $x, y \in E$ folgt

$$Q(x, y) = \frac{g(a, x, y)}{g(a, a, y)} = \frac{P(a, x)P(x, y)}{P(a, a)P(a, y)}.$$

Mit den Definitionen $q := \frac{1}{P(a, a)}$ und $r(x) := P(a, x)$ erhalten wir

$$Q(x, y)r(y) = qr(x)P(x, y). \quad (4.10)$$

4. Zusammenhänge von Markov-Feldern, Markov-Ketten und Gibbs-Maßen

Summieren wir über y und wird $\sum_y P(x, y) = 1$ berücksichtigt, so sieht man, dass q ein Eigenwert der positiven Matrix Q und r der dazugehörige Eigenvektor ist. Sei nun $q' \in \mathbb{C} \setminus \{q\}$ ein weiterer Eigenwert von Q mit Eigenvektor $r' \in \mathbb{C}^E$ dann gilt

$$\sum_y Q(x, y) r'(y) = q' r'(x).$$

Aus (4.10) folgt

$$Q(x, y) = q \frac{r(x)}{r(y)} P(x, y)$$

und damit

$$\sum_y q \frac{r(x)}{r(y)} P(x, y) r'(y) = q' r'(x).$$

Man erhält

$$\sum_y P(x, y) \frac{r'(y)}{r(y)} = \frac{q' r'(x)}{q r(x)}.$$

Wir haben also gezeigt, dass $\frac{q'}{q}$ Eigenwert der regulären Markov-Matrix P mit Eigenvektor $\left(\frac{r'(y)}{r(y)}\right)_{y \in E}$ ist. Da $q' \neq q$ und 1 der größte Eigenwert einer regulären Markov-Matrix ist, folgt $\left|\frac{q'}{q}\right| < 1$. Damit ist q der größte positive Eigenwert von Q und (4.5) definiert P . Da P durch Q und Q durch g , die γ -bestimmende Funktion, definiert wird, folgt die Injektivität von b .

Für den Beweis der Surjektivität von b benötigen wir den nächsten Beweisteil.

3.) Es sei $g : E^3 \rightarrow]0, \infty[$ die bestimmende Funktion von $\gamma \in \Gamma$ und $P \in \Pi$ werde durch g und ein $a \in E$ definiert. Wir zeigen nun, dass (4.9) äquivalent zu

$$\frac{g(x, y, z) g(a, x, a)}{g(x, a, z) g(a, a, a)} = \frac{g(a, x, y) g(a, y, z)}{g(a, a, y) g(a, a, z)} \quad \forall x, y, z \in E \quad (4.11)$$

ist. Mit (4.9) erhalten wir

$$\frac{g(x, y, z)}{g(x, a, z)} = \frac{P(x, y) P(y, z)}{P(x, a) P(a, z)} \quad \forall x, y, z \in E.$$

Nutzen wir (4.5) erhält man

$$\frac{g(x, y, z)}{g(x, a, z)} = \frac{Q(x, y) Q(y, z)}{Q(x, a) Q(a, z)}$$

und schließlich mit $Q(x, y) = \frac{g(a, x, y)}{g(a, a, y)}$

$$\frac{g(x, y, z)}{g(x, a, z)} = \frac{g(a, x, y) g(a, y, z) g(a, a, a) g(a, a, z)}{g(a, a, y) g(a, a, z) g(a, x, a) g(a, a, z)}$$

und damit das gewünschte Ergebnis.

4.) Wir zeigen, dass $b : \Pi \rightarrow \Gamma$ surjektiv ist.

4.3. Homogene Markov-Spezifizierungen und Gibbssche Spezifizierungen

Gegeben sei $\gamma \in \Gamma$ mit der bestimmenden Funktion g . Zu der Funktion erhalten wir mit einem $a \in E$ nach 2.) ein $P \in \Pi$. Wir müssen $\gamma = b(P)$ zeigen. Es sei $\tilde{\gamma} = b(P)$. Dann gilt $\gamma = \tilde{\gamma}$, wenn $\gamma_{\{i\}} = \tilde{\gamma}_{\{i\}}$ für alle $i \in S$. $\gamma_{\{i\}}$ ist durch (4.9) gegeben. Nach Teil 3 ist (4.9) äquivalent zu (4.11). Es bleibt also zu zeigen, dass die γ -bestimmende Funktion g Gleichung (4.11) erfüllt. Dazu setzen wir $\Lambda = \{1, 2\}$ und wählen ein beliebiges $z \in E$ fest. Mit der abkürzenden Schreibweise

$$[xy] := \gamma_{\Lambda}(\sigma_1 = x, \sigma_2 = y \mid \omega) \quad (x, y \in E),$$

wobei $\omega \in \Omega$ mit $\omega_0 = a$ und $\omega_3 = z$, erhalten wir für die Konsistenzbedingung $\gamma_{\Lambda} = \gamma_{\Lambda} \gamma_{\{1\}} = \gamma_{\Lambda} \gamma_{\{2\}}$ der Spezifizierung γ

$$[xy] = g(a, x, y) \sum_{u \in E} [uy] = g(x, y, z) \sum_{v \in E} [xv] \quad (x, y \in E).$$

Außerdem gilt

$$\begin{aligned} [ay] &= g(a, a, y) \sum_{u \in E} [uy] = g(a, y, z) \sum_{v \in E} [av] \\ [aa] &= g(a, a, z) \sum_{v \in E} [av]. \end{aligned}$$

Wir erhalten

$$\frac{[xy][ay]}{[ay][aa]} = \frac{[xy]}{[ay]} = \text{rechte Seite von (4.11)}.$$

Analog erhält man für die linke Seite von (4.11) ebenfalls das Verhältnis

$$\frac{[xy]}{[ay]}.$$

Damit erfüllt g Gleichung (4.11) und es gilt $\gamma = b(P)$.

5.) Abschließend ist zu zeigen, dass $\mathcal{G}(\gamma) \subset \{\mu_P\}$, wenn $P \in \Pi$ und $\gamma = b(P)$.

Es sei $i \in \mathbb{Z}$, $n \geq 0$, $\Lambda = \{i, i+1, \dots, i+n\}$, $k \geq 1$ und

$$\Delta = \Delta(k) = \{i - k + 1, \dots, i + n + k - 1\}.$$

Für $\zeta \in E^{\Lambda}$ und $\omega \in \Omega$ gilt

$$\begin{aligned} \gamma_{\Delta(k)}(\sigma_{\Lambda} = \zeta \mid \omega) &= \sum_{\eta \in E^{\Delta} : \eta_{\Lambda} = \zeta} \mu_P(\sigma_{\Delta} = \eta \mid \sigma_{\partial\Delta} = \omega_{\partial\Delta}) = \\ &= \frac{P^k(\omega_{i-k}, \zeta_i) P(\zeta_i, \zeta_{i+1}) \dots P(\zeta_{i+n-1}, \zeta_{i+n}) P^k(\zeta_{i+n}, \omega_{i+n+k})}{P^{n+2k}(\omega_{i-k}, \omega_{i+n+k})}. \end{aligned}$$

Der Ergodensatz liefert, dass unabhängig von $x \in E$

$$\lim_{k \rightarrow \infty} P^k(x, \cdot) = \alpha_P$$

4. Zusammenhänge von Markov-Feldern, Markov-Ketten und Gibbs-Maßen

gilt. Also bekommen wir

$$\lim_{k \rightarrow \infty} \gamma_{\Delta(k)}(\sigma_\Lambda = \zeta \mid \omega) = \alpha_P(\zeta_i) P(\zeta_i, \zeta_{i+1}) \dots P(\zeta_{i+n-1}, \zeta_{i+n}) = \mu_P(\sigma_\Lambda = \zeta).$$

Mit dem Satz von der beschränkten Konvergenz gilt für $\mu \in \mathcal{G}(\gamma)$

$$\mu(\sigma_\Lambda = \zeta) = \lim_{k \rightarrow \infty} \int \mu(d\omega) \gamma_{\Delta(k)}(\sigma_\Lambda \mid \omega) = \mu_P(\sigma_\Lambda = \zeta).$$

Damit gilt $\mu = \mu_P$. □

Jetzt können wir eine Charakterisierung der positiven homogenen Markov-Spezifizierung mit Hilfe der Gibbsschen Spezifizierung vornehmen.

Satz 4.3.3. *Eine Spezifizierung γ ist genau dann eine positive homogene Markov-Spezifizierung, wenn sie für ein homogenes Nächster-Nachbar-Potential eine Gibbssche ist.*

Beweis. γ sei eine positive homogene Markov-Spezifizierung. Der **Satz 4.3.2** liefert uns, dass wir die Spezifizierung mit Hilfe einer Markov-Matrix schreiben können. Ein Blick auf (4.6) liefert uns das Potential, vgl. **Definition 2.3.2**, für welches γ eine Gibbssche Spezifizierung ist

$$\Phi_A = \begin{cases} -\ln P(\sigma_i, \sigma_{i+1}) & \text{wenn } A = \{i, i+1\}, \\ 0 & \text{wenn } A = \{i\}. \end{cases} \quad (4.12)$$

Denn Einsetzen des Potentials in die Gibbsche Spezifizierung liefert wieder (4.6). Analog ist die Umkehrung zu beweisen. □

Bemerkung 4.3.3. Die Sätze 4.3.2 und 4.3.3 liefern mit den nötigen Voraussetzungen die Äquivalenz von Markov-Ketten und Gibbs-Feldern. Genauer kann gezeigt werden, dass $\mathcal{G}(\Phi) = \{\mu_P\}$ eine Bijektion definiert. Dafür müssten noch Definitionen und Sätze zur Äquivalenz von Potentialen gemacht werden, vgl. Kapitel 2.4 in [26].

4.4. Das eindimensionale Ising-Modell

Ein für die statistische Mechanik besonders wichtiges Beispiel für Markov-Ketten ist das eindimensionale Ising-Modell, vgl. [26]. Anfang der zwanziger Jahre des letzten Jahrhunderts beauftragte der Physiker Wilhelm Lenz (1888-1957) seinen Doktoranden Ernst Ising (1900-1998) mit der mathematischen Ausarbeitung des Modells eines eindimensionalen Ferromagneten. Ising promovierte 1924 mit der Schrift „Beitrag zur Theorie des Ferro- und Paramagnetismus“. Trotzdem wird es in der Literatur ausschließlich als Ising-Modell bezeichnet. Vereinfacht bestand Lenz’ und Isings Idee darin, den eindimensionalen ferromagnetischen Stoff durch \mathbb{Z} , das heißt durch die Menge der ganzen Zahlen zu beschreiben. Das Modell war der erste Schritt zur mathematischen Untersuchung von Phasenübergängen, auch wenn man erst mit den höherdimensionalen Ising-Modellen tatsächlich einen Phasenübergang beschreiben kann.

Wir wollen uns einen ferro- oder antiferromagnetischen Stoff als eindimensionales Gitter, also als eine Kette, von magnetischen Momenten vorstellen. Diese magnetischen Momente können die Zustände $+1$ und -1 bzgl. der y -Achse einnehmen. Für unsere angestrebte Markov-Kette bedeutet dies: $S = \mathbb{Z}$ und $E = \{-1, +1\}$. Mit der üblichen Notation erhalten wir den Konfigurationsraum $\Omega = E^S$ und den Zustand an der Stelle $i \in S$ als $\sigma_i \in \{-1, +1\}$. Zwischen nächsten Nachbarn gibt es eine Wechselwirkung, die wir mit dem Isingpotential Φ beschreiben:

$$\Phi_A = \begin{cases} -J\sigma_i\sigma_{i+1} & \text{für } A = \{i, i+1\} \\ -h\sigma_i & \text{für } A = \{i\} \\ 0 & \text{sonst} \end{cases}.$$

Dabei beschreibt J die Wechselwirkung und h ein mögliches äußeres Feld. Mit dem Potential können wir nun eine Gibbsche Spezifizierung γ^Φ bestimmen und damit definieren wir eine Menge von Gibbs-Maßen $\mathcal{G}(\gamma^\Phi)$. Um zu zeigen, dass der eindimensionale Fall keinen Phasenübergang liefert, müssen wir die Eindeutigkeit des Gibbs-Maßes beweisen, also $|\mathcal{G}(\gamma^\Phi)| = 1$. Die Gibbsche Spezifizierung ist nach **Satz 4.3.3** eine positive homogene Markov-Spezifizierung, die durch die erzeugende Funktion g

$$\begin{aligned} p(\sigma_i | \sigma_{i-1}, \sigma_{i+1}) &= g(\sigma_{i-1}, \sigma_i, \sigma_{i+1}) \\ &= \frac{\exp(\sigma_i(h + J\sigma_{i-1} + J\sigma_{i+1}))}{\exp((h + J\sigma_{i-1} + J\sigma_{i+1})) + \exp(-(h + J\sigma_{i-1} + J\sigma_{i+1}))}. \end{aligned}$$

gegeben ist. Nach **Satz 4.3.2** gilt $\mathcal{G}(\gamma^\Phi) = \{\mu_P\}$, wobei P die zur Markov-Spezifizierung gehörende Markov-Matrix bezeichnet. Mit Hilfe der stationären Verteilung der Markov-Kette können wir also das eindeutige Gibbs-Maß μ_P bestimmen. Ebenfalls mit **Satz 4.3.2** und $a = +1$ erhalten wir

$$Q = \begin{pmatrix} \frac{g(+, -, -)}{g(+, +, -)} & \frac{g(+, -, +)}{g(+, +, +)} \\ \frac{g(+, +, -)}{g(+, +, -)} & \frac{g(+, +, +)}{g(+, +, +)} \end{pmatrix} = \begin{pmatrix} \exp(-2h) & \exp(-2h - 4J) \\ 1 & 1 \end{pmatrix},$$

wobei $+$ = $+1$ und $-$ = -1 . Als größten positiven Eigenwert errechnet man mit dem charakteristischen Polynom

$$\det Q = (\exp(-2h) - \lambda)(1 - \lambda) - \exp(-2h - 4J) = 0$$

$$\lambda = \frac{1 + \exp(-2h)}{2} + \sqrt{\left(\frac{1 + \exp(-2h)}{2}\right)^2 - \exp(-2h) + \exp(-2h - 4J)}.$$

Ein dazugehöriger rechter Eigenvektor muss glücklicherweise nicht berechnet werden, da sich dieser für $\sigma_i = \sigma_j$ aus der Formel (4.5) herauskürzt. Des Weiteren gilt

$$\sum_{\sigma_i \in E} P(\sigma_j, \sigma_i) = 1.$$

4. Zusammenhänge von Markov-Feldern, Markov-Ketten und Gibbs-Maßen

Also erhalten wir die Markov-Matrix

$$P = \begin{pmatrix} P(+1, +1) & P(-1, +1) \\ P(+1, -1) & P(-1, -1) \end{pmatrix} = \begin{pmatrix} \lambda^{-1} & 1 - \exp(-2h)\lambda^{-1} \\ 1 - \lambda^{-1} & \exp(-2h)\lambda^{-1} \end{pmatrix}.$$

Die Matrix P ist ergodisch. Wir können also den Ergodensatz nutzen, um $\alpha_P(\sigma_i = +1)$ und $\alpha_P(\sigma_i = -1)$ zu berechnen. Beginnen wir mit dem Eigenvektor $v = (v_1, v_2)^T$ zum Eigenwert 1

$$\begin{array}{cc|c} \lambda^{-1} - 1 & 1 - \exp(-2h)\lambda^{-1} & 0 \\ 1 - \lambda^{-1} & \exp(-2h)\lambda^{-1} - 1 & 0 \\ \hline \lambda^{-1} - 1 & 1 - \exp(-2h)\lambda^{-1} & 0 \\ 0 & 0 & 0 \end{array}.$$

Setzen wir $v_1 = t \in \mathbb{R}$ erhalten wir für v_2

$$(\lambda^{-1} - 1)t + (1 - \exp(-2h)\lambda^{-1})v_2 = 0 \Leftrightarrow v_2 = \frac{-(\lambda^{-1} - 1)t}{1 - \exp(-2h)\lambda^{-1}}.$$

Gilt nun $v_1 + v_2 = 1$, so haben wir die gesuchte Wahrscheinlichkeitsverteilung $\alpha_P = (\alpha_P(\sigma_i = +1), \alpha_P(\sigma_i = -1))$. Damit folgt

$$v_1 + v_2 = 1 \Leftrightarrow t + \frac{(1 - \lambda^{-1})t}{1 - \exp(-2h)\lambda^{-1}} = 1$$

und

$$\begin{aligned} \alpha_P(\sigma_i = +1) &= \left(\frac{1 - \lambda^{-1}}{1 - \exp(-2h)\lambda^{-1}} + 1 \right)^{-1} = \frac{1 - \exp(-2h)\lambda^{-1}}{2 - \lambda^{-1} - \exp(-2h)\lambda^{-1}} \\ \alpha_P(\sigma_i = -1) &= 1 - \alpha_P(\sigma_i = +1). \end{aligned}$$

Mit $\cosh x = \frac{\exp(x) + \exp(-x)}{2}$ und $\sinh x = \frac{\exp(x) - \exp(-x)}{2}$ kommt man auf

$$\lambda = \exp(-h) \left(\cosh h + \sqrt{\exp(-4J) + \sinh^2 h} \right)$$

und

$$\alpha_P(\sigma_i = +1) = \frac{1}{2} \left(1 + \frac{\sinh h}{\sqrt{\exp(-4J) + \sinh^2 h}} \right).$$

Nun können wir z. B.: für $A = \{\sigma_i = +1, \sigma_{i+1} = +1, \sigma_{i+2} = -1\}$

$$\mu_P(A) = \alpha_P(\sigma_i = +1) \cdot P(+1, +1) \cdot P(+1, -1)$$

berechnen.

Wenn es keine Interaktion gibt, d. h. es gilt $J = 0$, vereinfacht sich λ zu

$$\lambda = 1 + \exp(-2h).$$

Für die Übergangsmatrix P folgt

$$P = \begin{pmatrix} (1 + \exp(-2h))^{-1} & 1 - \exp(-2h)(1 + \exp(-2h))^{-1} \\ 1 - (1 + \exp(-2h))^{-1} & \exp(-2h)(1 + \exp(-2h))^{-1} \end{pmatrix}$$

$$= \begin{pmatrix} (1 + \exp(-2h))^{-1} & (1 + \exp(-2h))^{-1} \\ \exp(-2h)(1 + \exp(-2h))^{-1} & \exp(-2h)(1 + \exp(-2h))^{-1} \end{pmatrix},$$

für $\alpha_P(\sigma_i = +1)$ und $\alpha_P(\sigma_i = -1)$ ergeben sich

$$\alpha_P(\sigma_i = +1) = \lambda^{-1} = (1 + \exp(-2h))^{-1},$$

$$\alpha_P(\sigma_i = -1) = 1 - \lambda^{-1} = \exp(-2h)(1 + \exp(-2h))^{-1}.$$

Damit vereinfacht sich die Berechnung von $\mu_P(A)$ zu

$$\mu_P(A) = \lambda^{-1} \lambda^{-1} \exp(-2h) \lambda^{-1}.$$

Es gilt

$$\alpha_P(\sigma_i = +1) = P(+1, +1) = P(-1, +1)$$

und

$$\alpha_P(\sigma_i = -1) = P(+1, -1) = P(-1, -1).$$

Die Ereignisse sind bei fehlender Interaktion, wie zu erwarten war, unabhängig.

4.5. Sequenzen von Nukleotiden

Eine weitere wichtige Anwendung von Gibbs-Maßen ist die Analyse von DNS (Desoxyribonukleinsäure). Die DNS ist ein Polynukleotid. Die einzelnen Nukleotide sind die vier Basen Adenin (A), Guanin (G), Cytosin (C) und Thymin (T). Abschnitte der DNS bilden die Gene. Diese Abschnitte legen durch die Reihenfolge der Nukleotide den Aufbau von Proteinen fest. Proteine bestehen aus bis zu 20 verschiedenen Aminosäuren. Jeweils drei Nukleotide, Codons, spezifizieren eine Aminosäure. Es gibt also 64 Codons, wobei drei Codons keine Aminosäure festlegen, sondern die Proteinsynthese beenden, die so genannten Stopcodons. Da es nur 20 Aminosäuren gibt, können unterschiedliche Codons die selbe Aminosäure definieren. Entscheidend ist also die Reihenfolge der Nukleotide. Und genau diese Reihenfolge kann mit Gibbs-Maßen modelliert werden.

Ich werde zwei Aspekte betrachten, zum einen die Nukleotid-Substitution und zum anderen das Auffinden von bestimmten Nukleotidsequenzen in der DNS, vgl. [37, 55, 77]. Beginne ich mit der Nukleotid-Substitution. Es kann spontan oder durch äußere Einflüsse zum Austausch von einzelnen Nukleotiden kommen, dies bezeichnet man als Nukleotid-Substitution. Wir stellen uns das Polynukleotid, die DNS, als Kette von n Nukleotiden vor und betrachten eine kontinuierliche Markov-Kette. Ein einfaches Modell ergibt sich, wenn alle Nukleotide unabhängig von einander betrachtet werden. Man gibt also die Übergangswahrscheinlichkeiten an, dass Basis 1 an der Stelle i der Sequenz zu Basis

4. Zusammenhänge von Markov-Feldern, Markov-Ketten und Gibbs-Maßen

2 wechselt. Statt einzelner Basen können auch Codons betrachtet werden. Dann ergibt sich eine Übergangsmatrix mit 64×64 Einträgen. Wir wollen ein etwas modifiziertes Modell betrachten, siehe [37]. Es sei E die Menge der Codons ohne die Stopcodons und $S = \{1, \dots, n\}$. Mit $z \in E^S$ wird der Zustand des Polynukleotids und mit $z_i \in E$ das Codon an der Stelle $i \in S$ beschrieben. An den Stellen $i = 0$ und $i = n + 1$ seien die Stopcodons fest gewählt. Alle anderen Codons z_i können sich beliebig entsprechend einer zu definierenden Sprungrate verändern. Nun kommt der für die Gibbs-Maße typische Aspekt ins Spiel, dass der Wechsel von z_i zu y_i von z_{i-1} und z_{i+1} abhängen darf und nicht nur von z_i . Wir bezeichnen analog zu Kapitel 3.5 mit

$$c(y_i; z_{i-1}, z_i, z_{i+1}) > 0$$

die Sprungrate. Sie gibt die Rate an, dass der Codon z_i zu y_i wechselt. Von Interesse ist nun natürlich das mögliche stationäre Maß der Markov-Kette. Ist dieses Maß ein Gibbs-Maß, so können wir nach den **Sätzen 4.3.2** und **4.3.3** Wahrscheinlichkeiten für bestimmte Sequenzen von Nukleotiden wieder mit einer Markov-Kette berechnen. Wir wünschen uns für das stationäre Maß also die Form

$$\mu(z) = \frac{1}{Z} \exp \left(\Phi_1(z_1) + \sum_{i=2}^n \Phi(z_{i-1}, z_i) + \Phi_n(z_n) \right). \quad (4.13)$$

Damit ist μ ein endliches Gibbs-Maß mit einem Nächster-Nachbar-Potential definiert durch $\Phi_1, \Phi_n : E \rightarrow \mathbb{R}$ und $\Phi : E \times E \rightarrow \mathbb{R}$ und der Normierungskonstanten Z . Der folgende Satz sagt uns, welche Voraussetzungen die Sprungrate c erfüllen muss, damit wir ein Gibbs-Maß als stationäres Maß erhalten, siehe [37]. Der Satz ergibt sich ebenfalls aus Kapitel 3.5.

Satz 4.5.1. *Gegeben sei ein endlicher Zustandsraum E . Gilt für die Sprungrate mit $c(y_i; z_{i-1}, z_i, z_{i+1}) > 0$, dass*

$$\frac{c(y_i; z_{i-1}, z_i, z_{i+1})}{c(z_i; z_{i-1}, y_i, z_{i+1})} = \begin{cases} \frac{\exp(\Phi(z_{i-1}, y_i) + \Phi(y_i, z_{i+1}))}{\exp(\Phi(z_{i-1}, z_i) + \Phi(z_i, z_{i+1}))}, & 1 < i < n \\ \frac{\exp(\Phi_1(y_1) + \Phi(y_1, z_2))}{\exp(\Phi_1(z_1) + \Phi(z_1, z_2))}, & i = 1 \\ \frac{\exp(\Phi(z_{n-1}, y_n) + \Phi(y_n, z_n))}{\exp(\Phi(z_{n-1}, z_n) + \Phi(z_n, z_n))}, & i = n \end{cases},$$

so ist das Gibbs-Maß (4.13) ein stationäres Maß der durch c definierten kontinuierlichen Markov-Kette. Das stationäre Maß ist eindeutig.

Beweis. Ich führe den Beweis für den Fall $1 < i < n$. Die Fälle $i = 1$ und $i = n$ verlaufen analog.

Der **Satz (3.4.1)** sagt uns, dass μ genau dann ein stationäres Maß der Markov-Kette mit Sprungrate c ist, wenn

$$\sum_{z \neq \tilde{z}} \mu(z) c(z, \tilde{z}) - \mu(\tilde{z}) c(\tilde{z}, z) = 0$$

gilt, wobei sich z und \tilde{z} nur durch ein Codon unterscheiden dürfen und c die Rate des Wechsels von $z \in E^S$ zu $\tilde{z} \in E^S$ ist. Unterscheiden sich z und \tilde{z} durch mehr als ein Codon, so ist $c = 0$. Ein reversibles Maß μ , d. h. es gilt

$$\mu(z)c(z, \tilde{z}) = \mu(\tilde{z})c(\tilde{z}, z) \quad \forall z \neq \tilde{z},$$

ist auch stationär. Wir zeigen, dass μ von der Form (4.13) reversibel und damit stationär ist. Es sei z gegeben und \tilde{z} unterscheide sich nur an der Stelle i von z mit $\tilde{z}_i = y_i$. Dann ist μ reversibel, wenn

$$\begin{aligned} & \frac{1}{Z} \exp \left(\Phi_1(z_1) + \sum_{i=2}^n \Phi(z_{i-1}, z_i) + \Phi_n(z_n) \right) c(y_i; z_{i-1}, z_i, z_{i+1}) \\ &= \frac{1}{Z} \exp \left(\Phi_1(\tilde{z}_1) + \sum_{i=2}^n \Phi(\tilde{z}_{i-1}, \tilde{z}_i) + \Phi_n(\tilde{z}_n) \right) c(z_i; z_{i-1}, y_i, z_{i+1}). \end{aligned}$$

Da $z_j = \tilde{z}_j$ für $j \neq i$ folgt mit der Voraussetzung,

$$\frac{c(y_i; z_{i-1}, z_i, z_{i+1})}{c(z_i; z_{i-1}, y_i, z_{i+1})} = \frac{\exp(\Phi(z_{i-1}, y_i) + \Phi(y_i, z_{i+1}))}{\exp(\Phi(z_{i-1}, z_i) + \Phi(z_i, z_{i+1}))},$$

dass μ der Form (4.13) ein reversibles Maß der kontinuierlichen Markov-Kette ist.

Da der Zustandsraum endlich ist beschreibt c nach Kapitel 3.3 die Markov-Kette eindeutig, da Kolmogorovs *backward equation* bzw. *forward equation* eindeutig lösbar sind. Die Markov-Kette ist irreduzibel, da $c(y_i; z_{i-1}, z_i, z_{i+1}) > 0$ oder genauer für zwei Zustände z und $\tilde{z} \in E^S$ gibt es eine endliche Folge von Zuständen $x_0 = z, x_1, \dots, x_n = \tilde{z}$ mit $c(x_{m-1}, x_m) > 0$ für $1 \leq m \leq n$, vgl. [16, 33]. Des Weiteren ist der Zustandsraum endlich, damit folgt aus **Bemerkung 3.4.4**, dass das stationäre Maß eindeutig bestimmt ist, vgl. [1, 7, 8, 16, 33]. \square

Jetzt gehen wir davon aus, dass wir als stationäres Maß ein Gibbs-Maß der Form (4.13) haben. Gesucht ist die Wahrscheinlichkeit, eine bestimmte Sequenz zu finden, eventuell unter der Bedingung, dass einige der Codons an einigen Stellen des Polynukleotids bekannt sind. Dafür müssen wir bei gegebener Sprungrate die Normierungskonstante des Maßes (4.13) bestimmen und nach Möglichkeit die Berechnung von bedingten Wahrscheinlichkeiten vereinfachen. Dabei werden uns die **Sätze 4.3.2** und **4.3.3** behilflich sein. Wir wissen nämlich, dass das endliche Gibbs-Maß äquivalent zu einer Markov-Kette ist. Das stationäre Maß, ein endliches Gibbs-Maß, einer stetigen Markov-Kette mit der Zeit als Parametermenge wird also in eine Markov-Kette mit diskreter Parametermenge verwandelt. Diesmal ist die Parametermenge die Position des Nukleotids. Jetzt definieren wir $Q(x, y) = \exp(\Phi(x, y))$ für $x, y \in E$. Damit ist $Q = (Q(x, y))_{x, y \in E}$ eine positive Matrix und der **Satz 4.1.2** anwendbar. Es sei q der PF-Eigenwert von Q und r der entsprechende rechte Eigenvektor mit positiven Einträgen. Dann gilt

$$\sum_{y \in E} Q(x, y)r(y) = qr(x), \quad \forall x \in E$$

4. Zusammenhänge von Markov-Feldern, Markov-Ketten und Gibbs-Maßen

und man erhält mit

$$T = (T(x, y))_{x, y \in E} = \left(\frac{Q(x, y)r(y)}{qr(x)} \right)_{x, y \in E}$$

eine Markov-Matrix. Lösen wir nun

$$\prod_{i=2}^n T(z_{i-1}, z_i) = \frac{r(z_n)}{q^{n-1}r(z_1)} \exp \left(\sum_{i=2}^n \Phi(z_{i-1}, z_i) \right)$$

nach $\exp(\sum_{i=2}^n \Phi(z_{i-1}, z_i))$ auf und setzen dies in (4.13) ein, bekommen wir

$$\mu(z) = \frac{1}{Z} \frac{q^{n-1}r(z_1)}{r(z_n)} \exp(\Phi_1(z_1) + \Phi_n(z_n)) \prod_{i=2}^n T(z_{i-1}, z_i)$$

und damit

$$Z = q^{n-1} \sum_{x, y \in E} \frac{r(x)}{r(y)} \exp(\Phi_1(x) + \Phi_n(y)) T^{n-1}(x, y).$$

Auch unser zweites Ziel, die Berechnung von bedingten Wahrscheinlichkeiten vereinfacht sich nun. Setzen wir $n = 6$ und berechnen die bedingte Wahrscheinlichkeit $\mu(\sigma_\Lambda = \zeta | \sigma_{\partial\Lambda} = \omega)$ mit $\Lambda = \{3, 4\} \in S$, $\sigma_\Lambda = \zeta = (z_3 = x_3, z_4 = x_4)$ und $\sigma_{\partial\Lambda} = \omega_{\partial\Lambda} = (z_1 = x_1, z_2 = x_2, z_5 = x_5, z_6 = x_6)$, wobei $x_i \in E$, so erhalten wir

$$\begin{aligned} \mu(\sigma_\Lambda = \zeta | \sigma_{\partial\Lambda} = \omega) &= \frac{\frac{q^{6-1}r(x_1)}{r(x_6)} \exp(\Phi_1(x_1) + \Phi_6(x_6)) \prod_{i=2}^6 T(x_{i-1}, x_i)}{\sum_{x_3, x_4 \in E} \frac{q^{6-1}r(x_1)}{r(x_6)} \exp(\Phi_1(x_1) + \Phi_6(x_6)) \prod_{i=2}^6 T(x_{i-1}, x_i)} \\ &= \frac{T(x_2, x_3)T(x_3, x_4)T(x_4, x_5)}{\sum_{x_3, x_4 \in E} T(x_2, x_3)T(x_3, x_4)T(x_4, x_5)} = \frac{T(x_2, x_3)T(x_3, x_4)T(x_4, x_5)}{T^3(x_2, x_5)}. \end{aligned}$$

Dabei meint T^3 die dritte Potenz der Markov-Matrix T . Im Vergleich dazu erhalten wir bei einer Berechnung der bedingten Wahrscheinlichkeit mit dem endlichen Gibbs-Maß (4.13)

$$\mu(\sigma_\Lambda = \zeta | \sigma_{\partial\Lambda} = \omega) = \frac{\exp(\Phi(x_2, x_3) + \Phi(x_3, x_4) + \Phi(x_4, x_5))}{\sum_{x_3, x_4 \in E} \exp(\Phi(x_2, x_3) + \Phi(x_3, x_4) + \Phi(x_4, x_5))}.$$

Gehen wir nun davon aus, dass wir bereits den zeitlich stationären Zustand der Nukleotid-Substitutionen vorliegen haben und statt der Sprungraten c die Übergangswahrscheinlichkeiten $T(z_{i-1}, z_i) > 0$ kennen. Die Matrix $T = (T(x, y))_{x, y \in E}$ ist eine positive Markov-Matrix und es gilt der **Satz 4.1.1**. Wir betrachten jetzt also eine Markov-Kette mit einem endlichen Zustandsraum und einer diskreten Parametermenge, die nicht die Zeit sondern die Position des Nukleotids beschreibt. Mit diesen Voraussetzungen werden wir wieder zur Darstellung mit Gibbs-Maßen kommen und versuchen, aus den Übergangswahrscheinlichkeiten ein Wahrscheinlichkeitsmaß zu konstruieren. Der erste Teil, die Darstellung als

bedingtes Gibbs-Maß gelingt mit den **Sätzen 4.3.2** und **4.3.3**. Danach gibt es eine einzu-eins Korrespondenz zwischen den positiven Markov-Matrizen und den homogenen Markov-Spezifizierungen, die wiederum äquivalent zu den Gibbsschen Spezifizierungen sind. Mit unseren jetzigen Voraussetzungen erhalten wir wie oben

$$\mu_T(\sigma_\Lambda = \zeta | \sigma_{\partial\Lambda} = \omega) = \frac{T(x_2, x_3)T(x_3, x_4)T(x_4, x_5)}{T^3(x_2, x_5)}.$$

Mit der Definition

$$\varphi_A = \begin{cases} -\ln T(z_{i-1}, z_i) & , \text{ wenn } A = \{i-1, i\} \\ 0 & , \text{ sonst} \end{cases}$$

bekommen wir als bedingtes Gibbs-Maß

$$\begin{aligned} \mu_\varphi(\sigma_\Lambda = \zeta | \sigma_{\partial\Lambda} = \omega) &= \frac{\exp\left(-\left(\varphi_{\{2,3\}}(x_2, x_3) + \varphi_{\{3,4\}}(x_3, x_4) + \varphi_{\{4,5\}}(x_4, x_5)\right)\right)}{\sum_{x_3, x_4 \in E} \exp\left(-\left(\varphi_{\{2,3\}}(x_2, x_3) + \varphi_{\{3,4\}}(x_3, x_4) + \varphi_{\{4,5\}}(x_4, x_5)\right)\right)} \\ &:= \frac{\exp(-H_\Lambda^\varphi(\zeta\omega))}{Z} := \frac{\exp(-\sum_{A \cap \Lambda \neq \emptyset} \varphi_A)}{Z}. \end{aligned}$$

Damit können wir bedingte Wahrscheinlichkeiten angeben. Uns fehlt allerdings das Wahrscheinlichkeitsmaß. Dabei hilft uns die Tatsache, dass wir eigentlich nicht an Polynukleotiden der Länge $n = 6$ interessiert sind, sondern an Polynukleotiden mit bis zu mehreren 10.000 Nukleotiden. Wie oben bereits erwähnt sind die Voraussetzungen des **Satzes 4.1.1** erfüllt. Also nähern wir das gesuchte Wahrscheinlichkeitsmaß an, indem wir als das fehlende Startmaß $p(i = 1)$, $i = 1 \in S$ für lange Polynukleotide die stationäre Verteilung α_T mit

$$\alpha_T = \lim_{i \rightarrow \infty} p(i) = \lim_{i \rightarrow \infty} T^i p(1), \quad i \in S$$

verwenden. Damit bekommen wir

$$\mu_T(z) \approx \alpha_T(z_1) \prod_{i=2}^n T(z_{i-1}, z_i) = \alpha_T(z_1) \exp\left(-\sum_{i=2}^n \varphi_{\{i-1, i\}}(z_{i-1}, z_i)\right).$$

Bemerkung 4.5.1. Wir haben hier nicht das Wissen, dass $T = (T(x, y))_{x, y \in E} = \left(\frac{Q(x, y)r(y)}{qr(x)}\right)_{x, y \in E}$ gilt, da wir direkt von den Übergangswahrscheinlichkeiten $T(x, y)$ ausgegangen sind. Mit den Informationen, die wir ausgehend von der Sprungrate haben, ergibt sich für das stationäre Maß

$$\alpha_T(z_i) = l(z_i)r(z_i), \quad z_i \in E,$$

wobei l der linke Eigenvektor von Q zum PF-Eigenwert q sei mit $\sum_{x \in E} l(x)r(x) = 1$. Da l der linke Eigenvektor von Q ist gilt

$$\sum_{x \in E} l(x)Q(x, y) = ql(y)$$

4. Zusammenhänge von Markov-Feldern, Markov-Ketten und Gibbs-Maßen

und damit

$$\sum_{x \in E} l(x)r(x) \frac{Q(x,y)r(y)}{qr(x)} = l(y)r(y)$$

bzw.

$$\sum_{x \in E} l(x)r(x)T(x,y) = l(y)r(y).$$

Also ist die Bedingung $T\alpha_T = \alpha_T$ für stationäre Verteilungen erfüllt.

5. Markov-Ketten in der Schule

In den letzten Jahren wurde der Umfang der Stochastik in der Schulmathematik deutlich ausgebaut, so dass auch die Markovketten für den Unterricht interessant geworden sind. Natürlich sind nicht alle Vorschläge dieses Kapitels in einem Kurs umsetzbar. Vielmehr sollen die Vorschläge als mögliche Ergänzung angesehen werden.

5.1. Einleitung

Mit dem Beschluss der Kultusministerkonferenz vom 24. 5. 2002 sind seit dem Abitur 2005 neue einheitliche Prüfungsanforderungen (EPA) für das Fach Mathematik anzuwenden, siehe [78]. Danach wird die allgemeinbildende Funktion des Mathematikunterrichts dadurch betont, dass folgende Grunderfahrungen ermöglicht werden:

1. Mathematik als formale Wissenschaft
2. Mathematik als anwendbare Wissenschaft
3. Mathematik als Mittel zur Ausbildung von heuristischen Fähigkeiten

Die Abiturprüfungen sind so zu gestalten, dass ein breites Spektrum von Kompetenzen an geeigneten Inhalten überprüft werden kann. Zu den Kompetenzen gehören u. a.:

1. angemessenes Verwenden von mathematischer Fachsprache
2. sachgerechtes, flexibles und kritisches Umgehen mit grundlegenden Begriffen, Sätzen, Verfahren und Algorithmen, auch zur Lösung innermathematischer Probleme
3. mathematisches Modellieren zur Lösung realitätsnaher Probleme
4. Verknüpfen von Inhalten aus verschiedenen mathematischen Themenbereichen

Bei den fachlichen Inhalten kommen den drei Gebieten Analysis, Lineare Algebra/Analytische Geometrie und Stochastik besondere Bedeutung zu. Im Bereich Lineare Algebra/Analytische Geometrie gibt es aufbauend auf dem Vektorbegriff drei alternative Inhaltsstränge:

- A1 vektorielle analytische Geometrie
- A2 Anwendungen von Matrizen bei Abbildungen
- A3 Anwendungen von Matrizen bei mehrstufigen Prozessen

5. Markov-Ketten in der Schule

Diese Angaben durch die EPA wurden im Oktober 2004 durch einen Erlass des niedersächsischen Kultusministeriums, die Rahmenrichtlinien, konkretisiert, vgl. [78]. Neben den Kernthemen der drei Gebiete Analysis, Lineare Algebra/Analytische Geometrie und Stochastik legen die Rahmenrichtlinien die möglichen Erweiterungen fest. Welche Erweiterung für den jeweiligen Abiturjahrgang zu wählen ist, wird vom Kultusministerium für jeden Jahrgang in der Regel drei Jahre vor der Abiturprüfung festgelegt. Markov-Ketten können bzw. müssen demnach im Bereich Lineare Algebra/Analytische Geometrie im Falle der Alternative A3, vgl. [70], und als Erweiterung im Bereich Stochastik, vgl. [71], unterrichtet werden. So können mit Markov-Ketten alle Grunderfahrungen vermittelt werden. Auch für die Aneignung der oben genannten Kompetenzen sind Markov-Ketten ideal. Man kann durch das Modellieren realitätsnaher Probleme sogar fächerübergreifenden unterrichten, wie schon die zwei Beispiele Ising-Modell und DNA zeigen, vgl. die **Abschnitte 4.4** und **4.5**, ergibt sich neben den Anwendungen in Wirtschaftswissenschaften eine Verbindung zu den Naturwissenschaften insbesondere zur Biologie und zur Physik. Und wie fruchtbar gerade das Zusammenspiel von Mathematik und Physik ist, wurde häufig betont, so z. B. von Courant in [9]. Nicht nur verschiedene Fächer lassen sich kombinieren, sondern es ergibt sich die großartige Möglichkeit, Analysis und Lineare Algebra mit der elementaren Wahrscheinlichkeitstheorie zu kombinieren. Schüler können das für das Verständnis von Mathematik wichtige Zusammenspiel verschiedener mathematischer Bereiche kennen lernen, siehe [70, 71]. So spielen Grenzwerte, Matrizen, Eigenvektoren und Eigenwerte eine zentrale Rolle bei der Bestimmung von stationären Wahrscheinlichkeitsverteilungen. Das Wissen der Schüler kann also wunderbar vernetzt werden. Des Weiteren ist ein sinnvoller Einsatz eines Computeralgebrasystems möglich, da Schüler nur selten die eigenhändige Eigenwert- und Eigenvektorberechnung beherrschen. Wenngleich Markov-Ketten hierfür eine Motivation bieten. Mit dem folgenden diskreten Putzer-Algorithmus zur Berechnung von P^n kann auch die Kompetenz des Umgehens mit Algorithmen (oben Kompetenz 2) gelehrt werden. Das Berechnen von hohen Potenzen der Matrix P mit dem Taschenrechner als Annäherung an den Grenzwert $\lim_{n \rightarrow \infty} P^n$ läßt sich umgehen und damit auch ein fataler Umgang mit dem Grenzwert in der Schule, selbst wenn das Matrizenkalkül zum Teil auf den Rechner verlagert wird. Auch der Umgang mit Summen wird motiviert. Der nicht-diskrete Putzer-Algorithmus dient der Berechnung von e^{At} und somit dem Lösen von Systemen von linearen gewöhnlichen Differentialgleichungen mit gegebenen Anfangswerten

$$y'(t) = Ay(t) + b(t), \quad y(t_0) = y_0, \quad t \in \mathbb{R}.$$

Dieser Aspekt verdeutlicht, dass die Markov-Ketten Systeme von Differenzengleichungen sind

$$y(n+1) = Ay(n), \quad y(n) = A^n y(0),$$

und damit die diskrete Analogie zu Systemen von linearen gewöhnlichen Differentialgleichungen. Differenzengleichungen sind ein mögliches Erweiterungsthema für das Abitur im Bereich Analysis, siehe Rahmenrichtlinien [78]. Beschäftigen wir uns also mit dem Putzer-Algorithmus.

5.2. Der diskrete Putzer-Algorithmus

Für den diskreten Putzer-Algorithmus braucht man den Satz von Cayley-Hamilton, siehe z. B. [22].

Satz 5.2.1. *Jede Matrix erfüllt ihre charakteristische Gleichung*

$$p(A) = \prod_{j=1}^N (A - \lambda_j I) = 0 \quad \text{oder} \quad A^N + a_1 A^{N-1} + \dots + a_{N-1} A + a_N I = 0.$$

Nach der Vorbereitung können wir jetzt den diskreten Putzer-Algorithmus entwickeln, vgl. [17]. Unser Ziel ist es, A^n in der Form

$$A^n = \sum_{j=1}^s u_j(n) M(j-1) \tag{5.1}$$

zu schreiben, wobei u_j eine noch zu bestimmende skalare Funktion ist und

$$M(j) = (A - \lambda_j I) M(j-1), \quad M(0) = I. \tag{5.2}$$

Es ergibt sich also iterativ

$$M(n) = \prod_{j=1}^n (A - \lambda_j I).$$

Mit dem Satz von Cayley-Hamilton folgt $M(n) = 0$, $\forall n \geq N$, so dass wir in Gleichung (5.1) das s durch ein N ersetzen können. Für $n = 0$ erhalten wir

$$A^0 = I = u_1(0)I + u_2(0)M(1) + \dots + u_N(0)M(N-1).$$

Diese Gleichung wird mit $u_1(0) = 1$ und $u_2(0) = u_3(0) = \dots = u_N(0) = 0$ erfüllt. Mit unserem Ziel, der Gleichung (5.1) folgt

$$\sum_{j=1}^N u_j(n+1)M(j-1) = AA^n = A \left(\sum_{j=1}^N u_j(n)M(j-1) \right) = \sum_{j=1}^N u_j(n)AM(j-1).$$

Mit (5.2) ergibt sich

$$\sum_{j=1}^N u_j(n+1)M(j-1) = \sum_{j=1}^N u_j(n) (M(j) + \lambda_j M(j-1)).$$

Jetzt müssen die Koeffizienten der Matrizen $M(j)$, $1 \leq j \leq N$, verglichen und die obigen Bedingungen für $u_j(0)$ berücksichtigt werden, so dass man auf die Gleichungen

$$u_1(n+1) = \lambda_1 u_1(n), \quad u_1(0) = 1$$

5. Markov-Ketten in der Schule

und

$$u_j(n+1) = \lambda_j u_j(n) + u_{j-1}(n), \quad u_j(0) = 0, \quad j = 2, 3, \dots, N$$

kommt. Als Lösungen dieser Gleichungen berechnet man

$$u_1(n) = \lambda_1^n, \quad u_j(n) = \sum_{i=0}^{n-1} \lambda_j^{n-1-i} u_{j-1}(i), \quad j = 2, 3, \dots, N \quad \text{für } n > 0.$$

Fassen wir den diskreten Putzer-Algorithmus für eine $N \times N$ -Matrix A mit den Eigenwerten λ_j zusammen

$$A^n = \sum_{j=1}^N u_j(n) M(j-1),$$

$$M(j) = (A - \lambda_j I) M(j-1), \quad M(0) = I,$$

$$u_1(n) = \lambda_1^n, \quad u_j(n) = \sum_{i=0}^{n-1} \lambda_j^{n-1-i} u_{j-1}(i), \quad j = 2, 3, \dots, N \quad \text{für } n > 0,$$

$$u_1(0) = 1 \quad \text{und} \quad u_j(0) = 0, \quad j = 2, 3, \dots, N.$$

Abschließend will ich mit Verweis auf [3] den bereits oben erwähnten Putzer-Algorithmus zur Berechnung von e^{tA} angeben.

Satz 5.2.2. *Es seien $\lambda_1, \dots, \lambda_N$ die Eigenwerte einer $N \times N$ -Matrix A . Wir definieren*

$$P_0(A) := I, \quad P_k(A) := \prod_{m=1}^k (A - \lambda_m I), \quad k = 1, 2, \dots, N.$$

Dann erhalten wir

$$e^{tA} = \sum_{k=0}^{N-1} r_{k+1}(t) P_k(A),$$

wobei

$$r_1(t) = e^{\lambda_1 t}$$

und

$$r_k(t) = e^{\lambda_k t} \int_0^t e^{-\lambda_k \tau} r_{k-1}(\tau) d\tau, \quad k = 2, \dots, N.$$

5.3. Beispiele

Es folgen einige Beispiele für den Unterricht in der gymnasialen Oberstufe. Auch Teile der Abschnitte 4.4 und 4.5 lassen sich für einen fächerübergreifenden Unterricht sowie für Arbeitsgemeinschaften nutzen. Den Schülern ist der **Satz 4.1.1** ohne Beweis bekannt, vgl. [32]. Es sei noch einmal betont, dass die in diesem Kapitel gemachten Vorschläge als Ergänzungen für die Schule angesehen und einzelne Vorschläge herausgesucht werden müssen. Im Normalfall ist der gesamte Stoff mit einer Klasse nicht zu bewältigen.

Beispiel 5.3.1. Berechnen Sie $x(n)$ mit dem diskreten Putzer-Algorithmus und der Startbedingung $x(0) = (1, 0, 0)^T$, so dass $x(n+1) = Ax(n)$, wobei

$$A = \begin{pmatrix} 1 & 0 & -1 \\ 0 & 1 & 0 \\ -1 & 0 & 1 \end{pmatrix}.$$

Es gilt

$$x(n) = A^n x(0).$$

Also müssen wir A^n mit dem Putzer-Algorithmus berechnen. Zunächst bestimmt man die Eigenwerte der Matrix A :

$$\det A = (1 - \lambda)^3 - 1 + \lambda = -\lambda(\lambda^2 - 3\lambda + 2) = 0$$

und damit

$$\lambda_1 = 0, \quad \lambda_2 = 1, \quad \lambda_3 = 2.$$

Für den Putzer-Algorithmus benötigen wir $M(0) = I$, $M(1)$, $M(2)$ und zur Kontrolle rechnen wir $M(3) = 0$ nach:

Es ergibt sich für $\lambda_1 = 0$ berechnet man

$$M(1) = (A - 0 \cdot I) \cdot M(0) = A = \begin{pmatrix} 1 & 0 & -1 \\ 0 & 1 & 0 \\ -1 & 0 & 1 \end{pmatrix},$$

für $\lambda_2 = 1$

$$M(2) = (A - 1 \cdot I) \cdot M(1) = \begin{pmatrix} 0 & 0 & -1 \\ 0 & 0 & 0 \\ -1 & 0 & 0 \end{pmatrix} \begin{pmatrix} 1 & 0 & -1 \\ 0 & 1 & 0 \\ -1 & 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 0 & -1 \\ 0 & 0 & 0 \\ -1 & 0 & 1 \end{pmatrix},$$

und für $\lambda_3 = 2$

$$M(3) = (A - 2 \cdot I) \cdot M(2) = \begin{pmatrix} -1 & 0 & -1 \\ 0 & -1 & 0 \\ -1 & 0 & -1 \end{pmatrix} \begin{pmatrix} 1 & 0 & -1 \\ 0 & 0 & 0 \\ -1 & 0 & 1 \end{pmatrix} = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}.$$

Es fehlen $u_i(n)$, $i = 1, 2, 3$:

$$u_1(0) = 1, \quad u_1(n) = 0^n = 0, \quad n > 0,$$

$$u_2(0) = 0, \quad u_2(n) = \sum_{i=0}^{n-1} 1^{n-1-i} u_1(i) = 1^{n-1} = 1,$$

$$u_3(0) = 0, \quad u_3(n) = \sum_{i=0}^{n-1} 2^{n-1-i} u_2(i) = \sum_{i=0}^{n-2} 2^i = \frac{1 - 2^{n-1}}{1 - 2} = 2^{n-1} - 1.$$

5. Markov-Ketten in der Schule

Damit erhalten wir

$$A^n = \begin{pmatrix} 1 & 0 & -1 \\ 0 & 1 & 0 \\ -1 & 0 & 1 \end{pmatrix} + (2^{n-1} - 1) \cdot \begin{pmatrix} 1 & 0 & -1 \\ 0 & 0 & 0 \\ -1 & 0 & 1 \end{pmatrix} = \begin{pmatrix} 2^{n-1} & 0 & -2^{n-1} \\ 0 & 1 & 0 \\ -2^{n-1} & 0 & 2^{n-1} \end{pmatrix}, \quad n > 0.$$

Das Ergebnis lässt sich für $n = 1, 2$ leicht überprüfen

$$A^2 = \begin{pmatrix} 2 & 0 & -2 \\ 0 & 1 & 0 \\ -2 & 0 & 2 \end{pmatrix}.$$

Als Lösung für die Differenzengleichung erhalten wir

$$x(n) = A^n x(0) = (2^{n-1}, 0, -2^{n-1})^T.$$

Die folgende Aufgabe ist eine Standardaufgabe, für deren Lösung der Ergodensatz genutzt werden kann, siehe [17].

Beispiel 5.3.2. Wir betrachten die Vererbung eines Merkmals bei Tieren. Dieses Merkmal wird durch ein Paar der Gene G und g eindeutig bestimmt. Die Tiere können also die Kombination (G, G) , (G, g) ($= (g, G)$) oder (g, g) besitzen. Jedes Tier erbt zufällig eines der Gene des Genpaares von jedem Elternteil. In jeder Generation gibt es die drei möglichen Zustände $s_1 = GG$, $s_2 = Gg$ und $s_3 = gg$. $p_i(n)$ soll die Wahrscheinlichkeit sein, dass s_i in der n -ten Generation auftritt. p_{ij} ist die Wahrscheinlichkeit, dass s_i in der $(n+1)$ -ten Generation auftritt, wenn s_j in der n -ten Generation vorliegt. Die Paarungen der Tiere vom Zustand s_1 bis s_3 finden nur mit Partnern des Zustands s_2 statt.

- Stellen Sie die Übergangsmatrix P mit

$$p(n+1) = P \cdot p(n) \text{ und } p(n) = (p_1(n), p_2(n), p_3(n))^T$$

auf.

- Ist die Übergangsmatrix ergodisch?
- Berechnen Sie $\lim_{n \rightarrow \infty} p(n)$.

Die Wahrscheinlichkeit p_{11} , dass bei der Paarung mit einem Tier vom Genpaar GG in der nächsten Generation wieder die Kombination GG auftritt, ist $1/2$. Von dem Elternteil mit der Kombination GG kommt mit der Wahrscheinlichkeit 1 das Gen G . Das Elternteil mit Gg vererbt mit der Wahrscheinlichkeit $1/2$ das Gen G . Bei Paarung mit einem Tier vom Genpaar Gg beträgt die Wahrscheinlichkeit p_{12} , dass in der nächsten Generation GG auftritt, $1/4$. Beide Elternteile vererben mit der Wahrscheinlichkeit $1/2$ das Gen G . Die Übergangsmatrix P lautet also

$$P = \begin{pmatrix} \frac{1}{2} & \frac{1}{4} & 0 \\ \frac{1}{2} & \frac{1}{2} & \frac{1}{2} \\ 0 & \frac{1}{4} & \frac{1}{2} \end{pmatrix}.$$

Es gilt $p_{ij} \geq 0$ und $\sum_{i=1}^3 p_{ij} = 1$ für $j = 1, 2, 3$. P ist also eine Markov-Matrix. Da

$$P^2 = \begin{pmatrix} \frac{3}{8} & \frac{2}{8} & \frac{1}{8} \\ \frac{4}{8} & \frac{4}{8} & \frac{4}{8} \\ \frac{1}{8} & \frac{2}{8} & \frac{3}{8} \end{pmatrix},$$

ist P eine ergodische Matrix. Um $\lim_{n \rightarrow \infty} p(n)$ zu berechnen, nutzen wir den Ergodensatz und bestimmen einen Eigenvektor v zum Eigenwert 1. Wir erhalten das Gleichungssystem

$$(P - 1 \cdot I) \cdot v = 0$$

bzw.

$$\begin{array}{ccc|c} -\frac{1}{2} & \frac{1}{4} & 0 & 0 \\ \frac{1}{2} & -\frac{1}{2} & \frac{1}{2} & 0 \\ 0 & \frac{1}{4} & -\frac{1}{2} & 0 \\ -\frac{1}{2} & \frac{1}{4} & 0 & 0 \\ 0 & -\frac{1}{4} & \frac{1}{2} & 0 \\ 0 & \frac{1}{4} & -\frac{1}{2} & 0 \end{array}.$$

Der Vektor $v = (1, 2, 1)^T$ ist ein Eigenvektor. Damit erhalten wir

$$\lim_{n \rightarrow \infty} p(n) = \frac{1}{1+2+1} \begin{pmatrix} 1 \\ 2 \\ 1 \end{pmatrix} = \begin{pmatrix} \frac{1}{4} \\ \frac{1}{2} \\ \frac{1}{4} \end{pmatrix}.$$

Nach einer leichten Abänderung der Aufgabe, sind die Voraussetzungen des Ergodensatzes nicht mehr erfüllt. Der Putzer-Algorithmus kann uns allerdings behilflich sein.

Beispiel 5.3.3. Wir betrachten die vorherige Aufgabe und verändern das Paarungsverhalten. Nun sollen die Paarungen der Tiere vom Zustand s_1 bis s_3 nur mit Partnern des Zustands s_1 bzw. s_3 geschehen.

- Stellen Sie für die zwei Fälle die Übergangsmatrizen P_1 und P_3 auf.
- Sind die Matrizen ergodisch?
- Existiert $\lim_{n \rightarrow \infty} P_i^n$, $i = 1, 3$?

Die Übergangsmatrix P_1 für Paarungen mit Tieren vom Zustand s_1 lautet

$$P_1 = \begin{pmatrix} 1 & \frac{1}{2} & 0 \\ 0 & \frac{1}{2} & 1 \\ 0 & 0 & 0 \end{pmatrix},$$

5. Markov-Ketten in der Schule

die Matrix P_3 für Paarungen mit Tieren vom Zustand s_3

$$P_3 = \begin{pmatrix} 0 & 0 & 0 \\ 1 & \frac{1}{2} & 0 \\ 0 & \frac{1}{2} & 1 \end{pmatrix}.$$

Beide Matrizen sind Markov Matrizen. Es gibt aber aufgrund der Nullzeilen kein $m \in \mathbb{N}$, so dass P_i^m positiv ist. Beide Matrizen sind also nicht ergodisch. Der Ergodensatz kann nicht genutzt werden.

Nun versuchen wir, mit dem Putzer-Algorithmus $\lim_{n \rightarrow \infty} P_i^n$, $i = 1, 3$ zu berechnen. Für P_1 bestimmen wir die Eigenwerte mit

$$\det(P_1 - \lambda \cdot I) = \det \begin{pmatrix} 1 - \lambda & \frac{1}{2} & 0 \\ 0 & \frac{1}{2} - \lambda & 1 \\ 0 & 0 & -\lambda \end{pmatrix} = (1 - \lambda) \left(\frac{1}{2} - \lambda \right) (-\lambda) = 0.$$

Also erhält man

$$\lambda_1 = 1, \lambda_2 = \frac{1}{2} \text{ und } \lambda_3 = 0.$$

Der nächste Schritt ist die Berechnung der Matrizen $M(i)$, $i = 1, 2, 3$

$$M(1) = (P_1 - \lambda_1 I) \cdot M(0) = \begin{pmatrix} 0 & \frac{1}{2} & 0 \\ 0 & -\frac{1}{2} & 1 \\ 0 & 0 & -1 \end{pmatrix},$$

$$M(2) = (P_1 - \lambda_2 I) \cdot M(1) = \begin{pmatrix} 0 & 0 & \frac{1}{2} \\ 0 & 0 & -1 \\ 0 & 0 & \frac{1}{2} \end{pmatrix}$$

und

$$M(3) = (P_1 - \lambda_3 I) \cdot M(2) = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}.$$

Es fehlen die Koeffizienten $u_i(n)$, $i = 1, 2, 3$

$$u_1(n) = 1, \quad u_2(n) = \sum_{i=0}^{n-1} \left(\frac{1}{2} \right)^{n-1-i} = \frac{1 - \left(\frac{1}{2} \right)^n}{1 - \frac{1}{2}} = 2 - 2 \cdot \left(\frac{1}{2} \right)^n, \quad u_3(n) = 2 - 2 \cdot \left(\frac{1}{2} \right)^{n-1}.$$

Da $\lambda_3 = 0$, verschwindet die Summe zur Berechnung von $u_3(n)$ mit Ausnahme vom Summanden für $i = n - 1$. Dann gilt $0^0 = 1$.

Jetzt können wir P_1^n angeben

$$P_1^n = 1 \cdot M(0) + \left(2 - 2 \cdot \left(\frac{1}{2} \right)^n \right) \cdot M(1) + \left(2 - 2 \cdot \left(\frac{1}{2} \right)^{n-1} \right) \cdot M(2) =$$

$$\begin{pmatrix} 1 & 1 - \left(\frac{1}{2}\right)^n & 1 - \left(\frac{1}{2}\right)^{n-1} \\ 0 & \left(\frac{1}{2}\right)^n & \left(\frac{1}{2}\right)^{n-1} \\ 0 & 0 & 0 \end{pmatrix}.$$

Damit ergibt sich der Grenzwert

$$\lim_{n \rightarrow \infty} P_1^n = \begin{pmatrix} 1 & 1 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}.$$

Da die Spalten von $\lim_{n \rightarrow \infty} P_1^n$ identisch sind, ist $\lim_{n \rightarrow \infty} p_1(n)$ unabhängig von $p_1(0)$ gleich

$$\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}.$$

Analog erhält man für P_3

$$\lim_{n \rightarrow \infty} P_3^n = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 1 & 1 & 1 \end{pmatrix}$$

und

$$\lim_{n \rightarrow \infty} p_3(n) = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}.$$

In beiden Fällen existieren also die Grenzwerte $\lim_{n \rightarrow \infty} p_i(n)$ und $\lim_{n \rightarrow \infty} P_i^n$, obwohl die Matrizen P_i nicht ergodisch sind.

Das nächste Beispiel demonstriert, dass es einen Vektor p mit $p = Ap$ für eine nicht ergodische Matrix geben kann, ohne dass die Grenzwerte $\lim_{n \rightarrow \infty} P^n$ und $\lim_{n \rightarrow \infty} p(n)$ existieren.

Beispiel 5.3.4. Gegeben sei die Matrix A mit

$$A = \begin{pmatrix} 0 & 1 & 1 \\ \frac{1}{2} & 0 & 0 \\ \frac{1}{2} & 0 & 0 \end{pmatrix}.$$

- Ist A eine Markov-Matrix oder sogar eine ergodische Matrix?
- Existiert $\lim_{n \rightarrow \infty} A^n$?
- Existiert ein Vektor p mit $p = Ap$?

5. Markov-Ketten in der Schule

A ist offensichtlich eine Markov-Matrix, jedoch nicht ergodisch. Wir berechnen

$$A^2 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \frac{1}{2} & \frac{1}{2} \\ 0 & \frac{1}{2} & \frac{1}{2} \end{pmatrix}.$$

Für A^3 bekommen wir wieder die Ausgangsmatrix A , so dass wir verallgemeinern können:

$$A^{2n} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \frac{1}{2} & \frac{1}{2} \\ 0 & \frac{1}{2} & \frac{1}{2} \end{pmatrix}, \quad n \in \mathbb{N}$$

und

$$A^{2n-1} = \begin{pmatrix} 0 & 1 & 1 \\ \frac{1}{2} & 0 & 0 \\ \frac{1}{2} & 0 & 0 \end{pmatrix}, \quad n \in \mathbb{N}.$$

Die Verallgemeinerung lässt sich als Übung leicht mit dem Putzer-Algorithmus nachrechnen. Allerdings ist der Aufwand an dieser Stelle vergleichsweise hoch. Die Eigenwerte von A lauten

$$\lambda_1 = 1, \quad \lambda_2 = -1 \quad \text{und} \quad \lambda_3 = 0.$$

Es ergeben sich die Matrizen

$$M(1) = \begin{pmatrix} -1 & 1 & 1 \\ \frac{1}{2} & -1 & 0 \\ \frac{1}{2} & 0 & -1 \end{pmatrix},$$

$$M(2) = \begin{pmatrix} 0 & 0 & 0 \\ 0 & -\frac{1}{2} & \frac{1}{2} \\ 0 & \frac{1}{2} & -\frac{1}{2} \end{pmatrix}$$

und die Faktoren

$$u_1(n) = 1, \quad u_2(n) = \begin{cases} 0 & , \quad n \text{ gerade} \\ 1 & , \quad n \text{ ungerade} \end{cases}, \quad u_3(n) = \begin{cases} 1 & , \quad n \text{ gerade} \\ 0 & , \quad n \text{ ungerade} \end{cases}.$$

Der Ergodensatz kann in diesem Beispiel nicht genutzt werden und im Gegensatz zur vorherigen Aufgabe existieren nach den obigen Ausführungen die Grenzwerte $\lim_{n \rightarrow \infty} A^n$ und $\lim_{n \rightarrow \infty} p(n)$ nicht. Dennoch existiert ein Vektor p mit $p = Ap$ und $\sum_i p_i = 1$:

$$p = \begin{pmatrix} \frac{1}{2} \\ \frac{1}{4} \\ \frac{1}{4} \end{pmatrix},$$

da

$$\begin{pmatrix} 0 & 1 & 1 \\ \frac{1}{2} & 0 & 0 \\ \frac{1}{2} & 0 & 0 \end{pmatrix} \cdot \begin{pmatrix} \frac{1}{2} \\ \frac{1}{4} \\ \frac{1}{4} \end{pmatrix} = \begin{pmatrix} \frac{1}{2} \\ \frac{1}{4} \\ \frac{1}{4} \end{pmatrix}.$$

6. Literaturverzeichnis

- [1] Anderson, W. J.: Continuous-Time Markov Chains. Springer-Verlag 1991
- [2] Adler, R. J.: The geometry of random fields. John Wiley & Sons 1981
- [3] Apostol, T. M.: Calculus Volume II. John Wiley & Sons 1969
- [4] Applebaum, D.: Probability and information. Cambridge University Press 1996
- [5] Bartlett, M. S.: An introduction to stochastic processes. Cambridge University Press 1960
- [6] Bauer, H.: Wahrscheinlichkeitstheorie und Grundzüge der Maßtheorie. 2. Auflage. De Gruyter 1974
- [7] Bremaud, P.: Markov Chains- Gibbs Fields, Monte Carlo Simulation and Queues. Springer 1999
- [8] Clarke, A. B.; Disney, R.L.: Probability and random processes: A first course with applications. 2. Auflage. John Wiley & Sons 1985
- [9] Courant, R.; Robbins, H.: Was ist Mathematik? Springer 1962
- [10] Dawson, D. A.: Synchronous and asynchronous reversible Markov systems. Canad. Math. Bull. Vol. 17, 1975, 633-649
- [11] Dmitri, Z.; Vladimir, M.; Gerd, R.: Statistical Mechanics of nonequilibrium processes. Akademie Verlag 1996
- [12] Dobrushin, R. L.: The description of a random field by means of conditional probabilities and conditions of its regularity. Theory of Probability and its Applications 13, 1968, 197-224
- [13] Dobrushin, R. L.: The problem of uniqueness of a Gibbs random field and the problem of phase transition. Functional Analysis and its Applications 2, 1968, 302-312
- [14] Dobrushin, R. L.: Gibbsian random fields. The general case. Functional Analysis and its Applications 3, 1969, 22-28
- [15] Durrett, R.: An introduction to infinite particle systems. Stochastic Processes and their Applications 11, 1981, 109-150

6. Literaturverzeichnis

- [16] Durrett, R.: Essentials of Stochastic Processes. Springer 1999
- [17] Elaydi, S. N.: An introduction to difference equations. Springer 1996
- [18] Elstrodt, J.: Maß - und Integrationstheorie. 3. Auflage. Springer 2002
- [19] Fernandez, R.; Maillard, G.: Chains with complete connections and one-dimensional Gibbs measures. *Electronic Journal of Probability* 9, 2004, 145-176
- [20] Fernandez, R.; Maillard, G.: Construction of a specification from its singleton part. Paper 05-288 auf www.ma.utexas.edu/mparc, 2005
- [21] Fernandez, R.: Gibbsianness and non-Gibbsianness in lattice random fields *in* Mathematical statistical physics, Session LXXXIII. Elsevier 2006
- [22] Fischer, Gerd: Lineare Algebra. Vieweg, 12. Auflage 2000
- [23] Gabriel, P.: Das Variationsprinzip für Gibbs-Maße. Dissertation 1995
- [24] Georges A., Le Doussal, P.: From equilibrium spin models to probabilistic cellular automata. *Journal of Statistical Physics* 54, 1989, 1011-1064
- [25] Georgii, H.-O.: Canonical Gibbs measures. Springer-Verlag 1979
- [26] Georgii, H.-O.: Gibbs Measures and Phase Transitions. De Gruyter 1988
- [27] Gibbs, J. W.: Elementare Grundlagen der statistischen Mechanik. Verlag von Johann Ambrosius Barth 1905
- [28] Gihman, I. I.; Skorohod, A.V.: The theory of stochastic processes I. Springer 1974
- [29] Gihman, I. I.; Skorohod, A.V.: The theory of stochastic processes II. Springer 1975
- [30] Goldstein, S.; Kuik, R.; Lebowitz J. L.; Maes C.: From PCAs to equilibrium systems and back. *Communications in Mathematical Physics* 125, 1989, 71-79
- [31] Greiner, W.; Neise, L.; Stöcker, H.: Thermodynamik und Statistische Mechanik. Verlag Harri Deutsch 1987
- [32] Griesel, H.; Postel, H.; Suhr, F.; Gundlach, A.: Elemente der Mathematik. Leistungskurs Stochastik. Schroedel Druck A 2007
- [33] Grimmet, G. R.; Stirzacker, D. R.: Probability and random processes. Oxford University Press 1990
- [34] Honerkamp, J.: Statistical physics. Springer 1998
- [35] Jänich, K.: Topologie. Springer-Verlag 1980
- [36] Jaynes, E. T.: Probability Theory. Cambridge University Press 2003

- [37] Jensen, J. L.; Pedersen, A.-M., K.: Probabilistic models of DNA sequence evolution with context dependent rates of substitution. *Advances in applied Probability* 32, 2000, 499-517
- [38] Karlin, S.; Taylor, H.M.: A first course in stochastic processes. 2. Auflage. Academic Press 1975
- [39] Kipnis, C.; Landim, C.: Scaling limits of interacting particle systems. Springer-Verlag 1999
- [40] Knauf, A.; Seiler, R.: Vorlesung Mathematische Physik II: Statistische Mechanik. 2004
- [41] Kozlov, O.; Vasilyev, N.: Reversible Markov chains with local interaction *in* Multicomponent random Systems. Dekker 1980
- [42] Krawetz, S., A.; Womble, D. D.: Introduction to Bioinformatics. Humana Press 2003
- [43] Künsch, H.: Time reversal and stationary Gibbs measures. *Stochastic Processes and their Applications* 17, 1984, 159-166
- [44] Lanford, O. E.; Ruelle, D.: Observables at infinity and states with short range correlations in statistical mechanics. *Communication in Mathematical Physics* 13, 1969, 194-215
- [45] Lavis, D. A.; Bell, G.M.: Statistical Mechanics of Lattice Systems I. Springer 1999
- [46] Lebowitz, J. L.; Montroll, E.W.: Nonequilibrium Phenomena I. North-Holland Publishing Company 1983
- [47] Lebowitz, J. L.; Montroll, E.W.: Nonequilibrium Phenomena II. North-Holland Publishing Company 1984
- [48] Lebowitz, J. L.: Simple models of equilibrium and nonequilibrium Phenomena. North-Holland Publishing Company 1987
- [49] Lebowitz, J. L.; Maes, C.; Speer E.R.: Statistical mechanics of probabilistic cellular automata. *Journal of Statistical Physics* 59, 1990, 117-170
- [50] Ligget, T. M.: Interacting particle systems. Springer-Verlag 1985
- [51] Meintrup, D.; Schäffler, S.: Stochastik. Springer 2005
- [52] Nolting, W.: Spezielle Relativitätstheorie, Thermodynamik. 4. Auflage. Vieweg 1999
- [53] Nolting, W.: Statistische Physik. 5. Auflage. Springer-Verlag 2005
- [54] Oxtoby, J. C.: Measure and category. 2. Auflage. Springer-Verlag 1980

6. Literaturverzeichnis

- [55] Pedersen, A.-M., K.; Wiuf, C.; Christiansen, F., B.: A codon-based model designed to describe lentiviral evolution. *Molecular Biology and Evolution* 15, 1998, 1069-1081
- [56] Pra P.D.; Louis, P.-Y.; Røelly, S.: Stationary measures and phase transition for a class of probabilistic cellular automata. *ESAIM: Probability and Statistics* 6, 2002, 89-104
- [57] Preston, C.: *Gibbs states on countable sets*. Cambridge University Press 1974
- [58] Preston, C.: *Random fields*. Springer-Verlag 1976
- [59] Prum, B.; Fort, J. C.: *Stochastic processes on a lattice and Gibbs measures*. Kluwer Academic Publishers 1991
- [60] Reed, M.; Simon, B.: *Methods of modern mathematical physics. I: Functional Analysis*. Academic press 1972
- [61] Rivet, J.-P.; Boon, J. P.: *Lattice Gas Hydrodynamics*. Cambridge University Press 2001
- [62] Ruelle, D.: *Statistical mechanics*. W.A. Benjamin, INC 1969
- [63] Ruelle, D.: *Thermodynamic Formalism*. Addison-Wesley Publishing Company 1978
- [64] Sherman, S.: Markov random fields and Gibbs random fields. *Israel Journal of Mathematics* 14, 1973, 92-103
- [65] Spitzer, F.: Markov random fields and Gibbs ensembles. *The American Mathematical Monthly* 78, 1971, 142-154
- [66] Spitzer, F.: *Random fields and interacting particle systems*. M.A.A. Summer Seminar, Mathematical Association of America 1971
- [67] Spitzer, F.: A variational characterization of finite Markov chains. *The Annals of Mathematical Statistics* 43, 1972, 303-307
- [68] Spitzer, F.: Phase Transition in one-dimensional nearest-neighbor systems. *Journal of Functional Analysis* 20, 1975, 240-255
- [69] Spohn, H.: *Large scale dynamics of interacting particles*. Springer-Verlag 1991
- [70] Tietze, U.-P.; Klika, M.; Wolpers, H.: *Mathematikunterricht in der Sekundarstufe II, Band 2*. Vieweg 2000
- [71] Tietze, U.-P.; Klika, M.; Wolpers, H.: *Mathematikunterricht in der Sekundarstufe II, Band 3*. Vieweg 2002
- [72] Toda, M.; Kubo, R.; Saito, N.: *Statistical physics I*. Springer-Verlag 1992

- [73] Tolman, R. C.: The principles of statistical mechanics. Dover Publications, Inc., 1980
- [74] Vasilyev, N. B.: Bernoulli and Markov stationary measures in discrete local interactions *in* Lecture Notes in Mathematics. 653, Springer-Verlag 1978
- [75] Wolf-Gladrow, D. A.: Lattice-Gas Cellular Automata and Lattice Boltzmann Models. Springer 2000
- [76] Wu, N.: The maximum entropy method. Springer 1997
- [77] Zharkikh, A.: Estimation of evolutionary distances between nucleotid sequences. Journal of Molecular Evolution 39, 1994, 315-329
- [78] <http://db2.nibis.de/1db/cuvo/ausgabe/index.php?mat1=7>

Lebenslauf

Name:	Tim Scharlau
Geburtsdatum/-ort:	7. Juli 1979 in Bremen
Familienstand:	ledig
Schulausbildung:	
1986 - 1990	Grundschule Pennigbüttel
1990 - 1992	Orientierungsstufe Osterholz-Scharmbeck
1992 - 1999	Gymnasium Osterholz-Scharmbeck
1999	Abitur
Zivildienst:	
August 1999 - Juli 2000	Deutsches Rotes Kreuz Osterholz
Studium:	
WS 2000/2001 - WS 2004/2005	Mathematik und Physik für das höhere Lehramt (TU Braunschweig)
Dezember 2004	1. Staatsexamen
wissenschaftliche Tätigkeit:	
1. 1. 2005 - 30. 4. 2008	wissenschaftlicher Mitarbeiter TU Braunschweig Institut Computational Mathematics AG Partielle Differentialgleichungen
Referendariat:	
seit 1. 5. 2008	Studienreferendar für das Lehramt am Gymnasium im Studienseminar Braunschweig